

**IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF SOUTH CAROLINA
COLUMBIA DIVISION**

IN RE BLACKBAUD, INC.,
CUSTOMER DATA BREACH
LITIGATION

Case No.: 3:20-mn-02972-JFA

MDL No. 2972

**MEMORANDUM OPINION AND
ORDER**

This matter is currently before the Court on Plaintiffs’ Motion for Class Certification (ECF No. 292). The motion has been fully briefed and is ripe for review. Each party has also moved to exclude portions of the reports and testimony of the others’ experts under Rule 702 of the Federal Rules of Evidence and *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, and those motions are fully briefed and pending as well. (ECF Nos. 335, 337, 339, 341, 343, 345, 418, 419, 420, 426). The Court held a three-day *Daubert* and class certification hearing from March 6–8, 2024 at which all pending motions were argued. Having reviewed Plaintiffs’ Amended Consolidated Class Action Complaint, the parties’ class certification briefs, the parties’ *Daubert* briefs, the expert’s reports, and the many volumes of exhibits, as well as the relevant law, this Court denies Plaintiffs’ Motion for Class Certification (ECF No. 292) because of Plaintiffs’ failure to meet their burden of proof as to ascertainability, grants in part Defendant’s Motion to Exclude the Report and Testimony of C. Matthew Curtin, CISSP (ECF No. 341), denies Plaintiffs’ Motion to Exclude the

Report and Testimony of Sonya Kwon (ECF No. 419), and denies as moot all other pending *Daubert* motions.¹

I. FACTUAL AND PROCEDURAL HISTORY

This case arises out of a data breach of Defendant Blackbaud's systems which occurred between approximately February 7, 2020 and May 20, 2020. Defendant is a publicly traded Software-as-a-Service ("SaaS") company incorporated in Delaware and headquartered in Charleston, South Carolina. (ECF No. 194, p. 86). The company provides data collection services to a wide variety of "social good entities" including arts and cultural organizations, faith communities, foundations, healthcare organizations, higher education institutions, individual change agents, K-12 schools, and nonprofit organizations. These entities make up Defendant's customers, and Defendant serves them by collecting and storing the Personally Identifiable Information ("PII") and Protected Health Information ("PHI") belonging to these customers' donors, patients, students, and congregants, which this Court will refer to as Defendant's "constituents." The constituents, rather than Defendant's customers, are the plaintiffs in this case. The parties estimate that as many as 1.5 billion constituents' data was exposed during the breach. (ECF No. 317-2, p. 75).

Although not directly pertinent to this order, the details of the data breach are as follows: Between February 7, 2020 and May 20, 2020, individuals this Court will refer to as "threat actors" infiltrated some of Defendant's data centers that are located in Massachusetts. (ECF No. 265). The threat actors' identity was and is unknown. The threat actors were able to access Defendant's remote desktop environment initially using a compromised customer account, and they were ultimately able to gain widespread access to Defendant's data centers. Plaintiffs allege that over

¹ Unredacted versions of these motions (ECF Nos. 293, 342, and 410) have also been filed on the record. The Court has cited to the unredacted filings throughout this Order for ease of reference.

400 terabytes of data was successfully exfiltrated, and the threat actors subsequently demanded that Defendant pay a ransom in exchange for their deletion of the data. Defendant paid the ransom, but it never received any proof that the data had been deleted. (ECF No. 293, p. 9). Plaintiffs argue that the breach was able to occur and remain undetected for months because Defendant did not have adequate safeguards in place to prevent the breach. (ECF No. 293, pp. 9–10). Plaintiffs also criticize Defendant’s remediation efforts after discovering the breach, contending that its response was negligent and misleading. (ECF No. 293, p. 9). Accordingly, Plaintiffs contend that putative class members’ data remains susceptible to misuse and is actively being marketed on the dark web. (ECF No. 293, pp. 10–11).

In total, approximately ninety thousand backup files belonging to thirteen thousand Blackbaud customers and containing data belonging to approximately 1.5 billion constituents were impacted by the breach.² (ECF No. 329, p. 13). As shown in the chart below, Defendant provides—or at one point provided—those customers with varying combinations of eleven separate products. (ECF No. 342, p. 9).

² The term “backup files” as used in this opinion refers to old customer data hosted by Defendant on the servers which were affected by the breach. The data on those servers included “copies of customer databases” as well as “loose files.” (ECF No. 317-2, p. 44, 61). This Court will use the term “customer backup files” to refer to the copies of customer databases at issue and the term “loose files” to refer to the other allegedly compromised data at issue.

Blackbaud Products			
Accounting Platforms	Fundraising Management	Marketing Campaign Management	Combined Fundraising / Marketing Management
Blackbaud Financial Edge NXT Accounting solution for gift entry processing that relates donor and financial information. Financial Edge Accounting platform for nonprofits.	Raisers Edge 7 Tracking for all aspects of fundraising. Blackbaud eTapestry Fundraising / donor management specifically designed as cost-effective solution for small nonprofits.	Blackbaud Direct Marketing Complex marketing campaign management solution integrated with Raiser's Edge or Blackbaud CRM.	Blackbaud CRM Enterprise-level fundraising, marketing, program management, specialized for major giving and high-volume marketing programs. Blackbaud Raiser's Edge NXT Fundraising / relationship management with integrated data, analytics, and marketing tools.
Education-Specific Solution	Arts and Culture-Specific Solution	Interactive Marketing Website Tool	Prospect Research
Education Edge Solution for educational institutions to manage attendance, academic records, schedules, and grading.	Blackbaud Altru Solution for arts and cultural organizations to manage admissions, membership, fundraising, merchandise, etc.	Blackbaud NetCommunity Online tool for building interactive websites, social networking, email marketing, online giving, and event registration.	ResearchPoint Solution for completing prospect research on current and potential donors.
ECF No. 333-20, Stio Decl. Ex. 78, at 9-10			

5

Defendant's customers can customize these products once they purchase them, and its customers have ultimate control over the data that is stored using these products, how it is stored, whether encrypted fields are used as designed by Defendant, and whether a product is customized to suit a given customer's specific needs.³ (ECF No. 329, p. 13). As a result of the data breach, nearly 90,000 backup files containing data belonging to the 13,000 aforementioned customers were accessed. In other words, the threat actors accessed a slew of customer backup files during the breach, as opposed to the "live" databases that Defendant also maintains. (ECF No. 293, p. 28; ECF No. 329, p. 13).

In this action, Plaintiffs represent a putative class of individuals (or "constituents") whose data was provided to Defendant's customers and was ultimately hosted by Defendant. They assert that their PII and PHI were compromised from February 7, 2020 to May 20, 2020, when threat

³ Encryption refers to a method of storing sensitive data wherein that data is "scrambled" or made unreadable to anyone except for the person or entity with a key to unscramble it. For this reason, encrypted data is not part of the data included in Plaintiffs' proposed class definitions, and it is also oftentimes exempted from the kinds of data that data privacy or consumer protection laws seek to specifically protect from exposure. *See* Section III.A.1.

actors successfully infiltrated Defendant's systems. After the breach was made public, lawsuits were filed in state and federal courts across the United States before eventually being consolidated into the instant multidistrict litigation ("MDL") case before this Court. The Initial Transfer Order placing the MDL in the District of South Carolina was entered on December 15, 2020 (ECF No. 1), and the Consolidated Amended Complaint containing all claims that survived Defendant's Motion to Dismiss (ECF No. 124) was filed on February 3, 2022. (ECF No. 194). Following this Court's Order Granting in Part and Denying in Part Plaintiffs' Motion to Stage Class Certification Briefing (ECF No. 285), the instant Motion to Certify a Class was filed, addressing the subset of claims that the parties were instructed to brief.⁴ (ECF No. 292).

Plaintiffs' Motion for Class Certification asks this Court to certify the following classes and sub-classes: "Nationwide negligence and gross negligence classes under Massachusetts common law" for "[a]ll natural persons residing in the United States whose unencrypted information was stored on the database of a customer identified in Exhibit A to Defendant's Revised Fact Sheet from February 7, 2020 to May 20, 2020"; a sub-class under the California Consumer Privacy Act ("CCPA") consisting of "[a]ll natural persons residing in California whose unencrypted information (1) was stored on the database of a customer identified in Exhibit A to Defendant's Revised Fact Sheet from February 7, 2020 to May 20, 2020 and (2) contains the combination of data elements identified in Appendix 2 to this memorandum"; a sub-class under the California Confidentiality of Medical Information Act ("CMIA") consisting of "[a]ll natural persons residing in California whose unencrypted information (1) was stored on the database of a customer identified in Exhibit A to Defendant's Revised Fact Sheet from February 7, 2020 to May 20, 2020 and (2) contains the combination of data elements identified in Appendix 2 to this

⁴ Plaintiffs' original Consolidated Amended Complaint asserted ninety causes of action. (ECF No. 77).

memorandum”; a sub-class under the New York General Business Law (“N.Y. GBL”) consisting of “[a]ll natural persons residing in New York (1) whose unencrypted information was stored on the database of a customer identified in Exhibit A to Defendant’s Revised Fact Sheet from February 7, 2020 to May 20, 2020, and (2) who viewed or were exposed to Blackbaud’s post-breach representations regarding the scope of the breach and the ‘confirmation’ of destruction by the cybercriminals”; and lastly a sub-class under the Florida Deceptive and Unfair Trade Practices ACT (“FDUTPA”) that seeks injunctive relief and would consist of “[a]ll natural persons residing in Florida (1) whose unencrypted information was stored on the database of a customer identified in Exhibit A to Defendant’s Revised Fact Sheet from February 7, 2020 to May 20, 2020 and (2) who viewed or were exposed to Blackbaud’s post-breach representations regarding the scope of the breach and the ‘confirmation’ of destruction by the cybercriminals.” (ECF No. 293, pp. 11–12). Plaintiffs ask this Court to certify these classes and sub-classes under Federal Rules of Civil Procedure 23(b)(2) and 23(b)(3). Defendant opposes Plaintiffs’ motion on numerous grounds. Chief among them is that Plaintiffs have failed to show that a class is ascertainable under Rule 23 and Fourth Circuit precedent and further that the basic elements of Rule 23(a)–(b) are not met because “individual issues predominate.” (ECF No. 329, pp. 11–12).

II. LEGAL STANDARD

A. Class Certification

1. Federal Rule of Civil Procedure 23

Federal Rule of Civil Procedure 23(a) provides that class certification is proper only if: “(1) the class is so numerous that joinder of all members is impracticable; (2) there are questions of law or fact common to the class; (3) the claims or defenses of the representative parties are

typical of the claims or defenses of the class; and (4) the representative parties will fairly and adequately protect the interests of the class.” Fed. R. Civ. P. 23(a).

Even if all elements of Rule 23(a) are met, the proposed classes and sub-classes must satisfy one of the three additional requirements for certification found in Rule 23(b). *See EQT Prod. Co. v. Adair*, 764 F.3d 347, 357 (4th Cir. 2014) (quoting *Gunnells v. Healthplan Servs., Inc.*, 348 F.3d 417, 423 (4th Cir. 2003)). Plaintiffs in this case seek certification under Rules 23(b)(2) and 23(b)(3). (ECF No. 293, p. 11). Thus, as to all classes and sub-classes except for the proposed FDUTPA Sub-class, Plaintiffs must show that “questions of law or fact common to class members predominate over any questions affecting only individual members, and that a class action is superior to the other available methods of fairly and efficiently adjudicating the controversy.” Fed. R. Civ. P. 23(b)(3). “The predominance requirement is similar to but ‘more stringent’ than the commonality requirement of Rule 23(a).” *Thorn v. Jefferson-Pilot Life Ins. Co.*, 445 F.3d 311, 319 (4th Cir. 2006) (citing *Lienhart v. Dryvit Sys.*, 255 F.3d 138, 146 n.4 (4th Cir. 2001)). With respect to the proposed FDUTPA sub-class, which seeks injunctive relief, Plaintiffs must show that “the party opposing the class has acted or refused to act on grounds that apply generally to the class, so that final injunctive relief or corresponding declaratory relief is appropriate respecting the class as a whole.” Fed. R. Civ. P. 23(b)(2).

A party must produce enough evidence to demonstrate that class certification is in fact warranted. *See Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 350 (2011). If one of the requirements necessary for class certification is not met, the effort to certify a class must fail. *See Clark v. Experian Information Solutions, Inc.*, 2001 WL 1946329, at *4 (D.S.C. Mar. 19, 2001) (citing *Harriston v. Chicago Tribune Co.*, 992 F.2d 697, 205 (7th Cir. 1993)). The court must go beyond the pleadings, take a “‘close look’ at relevant matters,” conduct “a ‘rigorous analysis’ of

such matters,” and make “‘findings’ that the requirements of Rule 23 have been satisfied.” *See Gariety v. Grant Thornton, LLP*, 368 F.3d 356, 365 (4th Cir. 2004) (cleaned up). While the court should not “include consideration of whether the proposed class is likely to prevail ultimately on the merits,” *id.* at 366 (citing *Eisen v. Carlisle and Jacquelin*, 417 U.S. 156, 177–78 (1974)), “sometimes it may be necessary for the district court to probe behind the pleadings before coming to rest on the certification question.” *Id.* (citing *Gen. Tel. Co. of the Southwest v. Falcon*, 457 U.S. 247, 160 (1982)).

2. Ascertainability

The ascertainability requirement is a judicially-imposed gloss on Federal Rule of Civil Procedure 23 that has been “repeatedly recognized” by the Fourth Circuit and which requires that “the members of a proposed class [must] be ‘readily identifiable.’” *In re Marriott International, Inc., Customer Data Sec. Breach Litig.*, 341 F.R.D. 128, 143 (D. Md. 2022), *vacated and remanded sub nom. In re Marriott Int’l, Inc.*, 78 F.4th 677 (4th Cir. 2023), and *reinstated by In re Marriott Int’l Customer Data Sec. Breach Litig.*, 345 F.R.D. 137 (D. Md. 2023). This threshold ascertainability requirement is two-pronged. *Id.* “First, a class cannot be certified unless a court can readily identify the class members in reference to objective criteria.” *Id.* (quoting *EQT Prod. Co.*, 764 F.3d 347, 358) (4th Cir. 2014)) (cleaned up). “Second, there must be an administratively feasible way for the court to determine whether a particular individual is a class member.” *Id.* (quoting *Krakauer*, 925 F.3d 643, 658) (4th Cir. 2019)) (cleaned up); *see also EQT Prod. Co.*, 764 F.3d at 358 (stating that class certification is inappropriate when “class members are impossible to identify without extensive and individualized fact-finding”).

Courts in this circuit have used differing approaches to determine whether a proposed method of ascertaining a class is “administratively feasible.” *In re Marriott*, 341 F.R.D. at 144.

One approach holds that “the individualized fact-finding giving rise to mini-trials that defeat ascertainability are those requiring determinations on the merits—not an administrative review to determine whether an objective element of a class definition is met.” *Id.* (cleaned up). Another approach considers whether “exceptionally complicated administrative review” is required and holds that sufficiently complicated administrative review can preclude ascertainability. *Id.* (citing *Spotswood v. Hertz Corp.*, No. 16-1200, 2019 WL 498822, at *6–8 (D. Md. Feb. 7, 2019)). The Fourth Circuit, in *Career Counseling, Inc. v. AmeriFactors Fin. Grp., LLC*, has recently intimated that the latter view—that exceptionally complicated administrative review can preclude ascertainability—is the more appropriate standard. The phrase “administratively feasible” implies that courts are permitted to consider the facts of the case before them and find that a certain degree of administrative review cannot reasonably be undertaken, using the methods and resources that the court and the parties have at their disposal. *See Career Counseling, Inc. v. AmeriFactors Fin. Grp., LLC*, No. 3:16-CV-03013-JMC, 2021 WL 3022677 (D.S.C. July 16, 2021), *aff’d*, 91 F.4th 202 (4th Cir. 2024) (stating that a class could not be ascertained due to the need to make at least 20,000 individualized inquiries “to determine if the fax number identified in the fax log [] was linked to a stand-alone fax machine on June 28, 2016”)). However, the Court’s ultimate conclusion would remain unchanged under either standard.

B. *Daubert* and Rule 702

Importantly, “two approaches . . . have emerged in the case law” with respect to making *Daubert* decisions at the class certification stage of a class action lawsuit. 3 William B. Rubenstein, et al., *Newberg on Class Actions* § 7:24 (5th ed. 2021). Courts are split between a “limited, focused, and perhaps tentative” application of *Daubert* and engaging in “a full and conclusive *Daubert* analysis . . . and assessment of the expert’s persuasiveness.” *Id.* The Fourth Circuit has yet to rule

on whether a full *Daubert* analysis at the class certification stage is always appropriate, appropriate in some circumstances, or is never appropriate. *Id.* “However, reported decisions suggest that courts in the . . . Fourth . . . Circuit[] follow a serious *Daubert* approach . . .” *Id.* This court is persuaded by the same rationale that has persuaded other courts in this circuit and which was carefully laid out by the Seventh Circuit in *American Honda Motor Co. v. Allen*, which held that:

[W]hen an expert’s report or testimony is critical to class certification . . . a district court must conclusively rule on any challenge to the expert’s qualifications or submissions prior to ruling on a class certification motion. That is, the district court must perform a full *Daubert* analysis before certifying the class if the situation warrants. . . . [T]he court must also resolve any challenge to the reliability of information provided by an expert if that information is relevant to establishing any of the Rule 23 requirements for class certification.

Am. Honda Motor Co. v. Allen, 600 F.3d 813, 815–16 (7th Cir. 2010); *see also* 3 William

B. Rubenstein et al., *Newberg on Class Actions* § 7:24 (5th ed. 2021).

Federal Rule of Evidence 702 states that:

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if the proponent demonstrates to the court that it is more likely than not that:

- (a) the expert’s scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or determine a fact in issue;
- (b) the testimony is based on sufficient facts or data;
- (c) the testimony is the product of reliable principles and methods; and
- (d) the expert’s opinion reflects a reliable application of the principles and methods to the facts of the case.

Fed R. Evid. 702 (2011).⁵

⁵ Rule 702 was amended in December of 2023. The parties, in much of their briefing on the motions addressed below, used the previous version of Rule 702. However, this Court’s analysis is not changed by the amendments. The amendment of the first sentence of Rule 702 only emphasizes the standard that a proponent of expert testimony must meet for that testimony to be admissible, which had been stated previously by many courts, including the United States Supreme Court. *See* Fed. R. Evid. 702 Advisory Committee’s Note to 2023 amendments. The amendment of Rule 702(d) “emphasize[d] that each expert opinion must stay within the bounds of what can be concluded from a reliable application of the expert’s basis and methodology.” *Id.*

District courts must serve as “gatekeepers to exclude unreliable expert testimony.”⁶ Fed. R. Evid. 702 2000 Advisory Committee Notes. The Supreme Court has noted that Federal Rule of Evidence 702 “imposes a special obligation upon a trial judge to ‘ensure that any and all scientific testimony . . . is not only relevant, but reliable.’” *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 589 (1993). “This entails a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid, *Id.* at 592–93, and whether the expert has “faithfully appl[ied] the methodology to the facts.” *Roche v. Lincoln Prop. Co.*, 175 Fed.Appx. 597, 602 (4th Cir. 2006). To determine whether the expert’s testimony is “scientifically valid,” courts have considered a variety of factors. *Daubert*, 509 U.S. at 593. Those factors include: (1) whether the theory or technique in question can or has been tested; (2) whether the theory or technique has been subjected to peer review and publication; (3) the known or potential rate of error; and (4) whether the theory or technique is generally accepted in the “relevant scientific community.” *Id.* at 593–94. As a court considers these factors, it must “focus . . . solely on principles and methodology,” not on an expert’s ultimate conclusions. *Id.* at 595. Notably,

⁶ Judge Paul W. Grimm, in a thorough and well-written *Daubert* order issued in the still-pending case *In re Marriott Int’l, Inc., Customer Data Security Breach Litigation*, includes a helpful discussion of what experts and parties should do to enable judges to most effectively perform their required gatekeeper function. Judge Grimm specifically notes that “it would be wise for [expert reports] to be written with the recognition that the trial judge is, in essence, an audience of one who must understand the opinions that will be testified to (and, vitally, why they are reliable), and how the methodology has been applied to the facts of the particular case.” *In re Marriott Int’l, Inc., Customer Data Sec. Breach Litig.*, 602 F. Supp. 3d 767, 774–75 (D. Md. 2022). Further, Judge Grimm states that “it would be of enormous help for the expert and the parties to provide the judge with copies of the most important peer expert articles . . . [with] key passages highlighted to facilitate review by the judge” and that it would be useful “if [counsel and experts] organized their memoranda to address the key *Daubert* factors and the evidence supporting their position with respect to each.” *Id.* Lastly, Grimm notes that expert reports are far too often “drafted as if their intended audience was another expert in the same field” and are “filled with undefined technical jargon and calculations likely to be undecipherable to a generalist judge; key peer reviewed literature is neither clearly identified, nor attached as an exhibit . . . and counsel fail to organize their memoranda in a manner that allows the judge to undertake the required analysis.” *Id.* Counsel in this case are encouraged to review Judge Grimm’s opinion and consider the guidance provided by Judge Grimm both for purposes of this case and for the benefit of other judges they will appear before in the future.

Daubert's list of factors "[is] meant to be helpful, not definitive." *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 151 (1999).

In addition to the factors listed in *Daubert*, courts have looked to additional factors to evaluate whether an expert's testimony may be considered by the trier of fact. These factors include: (1) whether an expert is "proposing to testify about matters growing naturally and directly out of research they have conducted independent of litigation, or whether they have developed their opinions expressly for the purposes of testifying,"; (2) whether the expert has unjustifiably extrapolated from an accepted premise to an unfounded conclusion; (3) whether the expert has considered alternative explanations for their conclusions; (4) whether the expert "is being as careful as he would be in his regular professional work outside his paid litigation consulting"; and (5) whether the expert's "discipline itself lacks reliability." See *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 43 F.3d 1311, 1317 (9th Cir. 1995); see also *General Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997) (noting that "[a] court may conclude that there is simply too great an analytical gap between the data and the opinion proffered"); *Claar v. Burlington N.R. Co.*, 29 F.3d 499, 502–03 (9th Cir. 1994); *Sheehan v. Daily Racing Form, Inc.*, 104 F.3d 940, 942 (7th Cir. 1997); *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 151 (1999); Daniel J. Capra, Stephen A. Saltzburg & Christine M. Arguello, *Evidence: The Objection Method* 532–33 (6th ed. 2021). These factors are useful tools in analyzing whether an expert's testimony is reliable, but courts also must not "overlook the evidentiary forest for the many scientific and technical trees." *In re Marriott Int'l* at 774. "There are four related and sometimes overlapping concepts that help guide a trial judge in deciding a *Daubert* challenge . . . [whether the evidence is] relevant . . . reliable . . . helpful . . ." and whether the evidence "fit[s] the facts and issues of the specific case." *Id.* (citing *Daubert*, 509 U.S. at 591–93).

Ultimately, “the proponent of the testimony must establish its admissibility by a preponderance of proof.” *Cooper v. Smith & Nephew, Inc.*, 259 F.3d 194, 199 (4th Cir. 2001) “[T]he trial court’s role as a gatekeeper is not intended to serve as a replacement for the adversary system, and consequently, the rejection of expert testimony is the exception rather than the rule.” *United States v. Stanley*, 533 F. App’x 325, 327 (4th Cir. 2013) (quoting Fed. R. Evid. 702 advisory committee’s note). Conversely, while Rule 702 was intended to liberalize the introduction of relevant evidence, courts “must recognize that due to the difficulty of evaluating their testimony, expert witnesses have the potential to ‘be both powerful and quite misleading.’” *Westberry v. Gislaved Gummi AB*, 178 F.3d 257, 261 (4th Cir. 1999) (quoting *Daubert*, 509 U.S. at 595). Thus, a trial court must balance the aim of the *Daubert* court, which was to ensure that trial courts serve their gatekeeper function, with the understanding that trial courts should not prevent the adversarial system from performing its proper function. *See Stanley*, 533 F. App’x at 327; *see also Daubert*, 509 U.S. at 589–90.

The Fourth Circuit has emphasized in recent years the importance of the gatekeeping role that trial judges are to play when undertaking a *Daubert* analysis. *See Sardis v. Overhead Door Corp.*, 10 F.4th 268, 290 (4th Cir. 2021) (stating that “[t]he trial court’s gatekeeping function requires more than simply taking the expert’s word for it”); *see also United States v. Fultz*, 590 F. App’x 226, 226 (4th Cir. 2015) (stating that a “careful analysis into reliability” is “crucial . . . [b]ecause ‘expert witnesses have the potential to be both powerful and quite misleading’”). This Court takes its gatekeeping role seriously, and its analysis of the expert opinions submitted by both parties is undertaken with the Fourth Circuit’s recent re-emphasis of that gatekeeping role in mind.

III. ANALYSIS

A. Plaintiffs' Motion for Class Certification

Plaintiffs have moved this Court under Rules 23(a), 23(b)(2), and 23(b)(3) of the Federal Rules of Civil Procedure to certify numerous classes and sub-classes as listed above. Each class and sub-class definition includes the requirement that a putative plaintiff's "unencrypted information was stored on the database of a customer identified in Exhibit A to Defendant's Revised Fact Sheet from February 7, 2020 to May 20, 2020," in addition to other state law-specific requirements for the proposed state law sub-classes. (ECF No. 293, pp. 11–12). Plaintiffs argue that they have satisfied Rule 23(a)'s requirements because: (1) millions of individuals' data was exposed in the breach, satisfying the numerosity requirement; (2) many common questions of law and fact are raised regarding Blackbaud's data management and security practices, the materiality of its post-breach representations, and the damages incurred by putative class members, satisfying the commonality requirement; (3) Plaintiffs' claims are all "premised on the same course of conduct and the same legal theory as all other class members," because they all arise out of the same data breach and set of alleged harms, thereby satisfying the typicality requirement; and (4) there are no conflicts between Class Counsel and class members, Plaintiffs have thus far prosecuted the action vigorously, and the class representatives' alleged harms and requested damages are the same as those suffered by other putative class members, thereby satisfying the adequacy requirement. (ECF No. 293, pp. 23–26).

Plaintiffs further assert that they satisfy the requirements of Rule 23(b)(2)–(3) because common questions of both law and fact predominate over individual issues regarding the breach, particularly in light of this Court's ruling that one state's law applies to the Plaintiffs' negligence claims. (ECF No. 293, pp. 31–32). Plaintiffs also argue that class treatment is the superior method

of litigating this case because of the small value of the individual claims, the existence of numerous lawsuits against Defendant that have been consolidated into the instant MDL, and the abundance of common issues that predominate in this matter together with the difficulty that individual litigation of the underlying claims would impose on the judicial systems in this country. (ECF No. 293, pp. 51–53).

Defendant argues that Plaintiffs’ case must fail as a threshold matter because Plaintiffs are unable to ascertain the class of individuals whose data was stored in Defendant’s backup files without extensive and individual fact-finding, nor have they shown that they can identify the affected data elements belonging to those individuals as required by their class and sub-class definitions. (ECF No. 329, pp. 15–17). Instead, Defendant argues that Matthew Curtin, Plaintiffs’ expert on ascertainability (“Curtin”) has presented an unreliable method for ascertaining putative plaintiffs and their data. Defendant asserts that Curtin’s method would require a prohibitively large and exhaustive restoration, organization, and search of the 90,000 backup files for each putative plaintiff and their respective data elements, in addition to the development of “data cleansing and standardization” and “validation” processes. (ECF No. 329, p. 23). Individualized inquiries with respect to each putative plaintiff would be required to determine, at minimum, (1) whether a putative plaintiff was a constituent of a customer whose backup files were accessed during the breach at the time of the breach, (2) which pieces of their PII or PHI were exposed, if any. Defendant further asserts that the lack of administrative feasibility inherent in Curtin’s proposed method of ascertaining a class is magnified when it is applied to the proposed sub-classes, each of which requires the Plaintiffs to prove state of residency and to identify specific types of data or lists of data elements per plaintiff under each state law in question. (ECF No. 329, pp. 24–27). Defendant also contends that Plaintiffs have failed to satisfy Rule 23’s commonality,

predominance, and typicality requirements because their damages and causation arguments would require “considerable individual inquiry,” and the class representatives “face the prospect of unique defenses.” (ECF No. 329, pp. 28–29). Lastly, Defendant argues that Plaintiffs lack standing because their damages theories do not present a cognizable injury under Article III because mere access of data does not constitute an injury in fact, some of the accessed data was already made public by Plaintiffs themselves, and the “lost market value” damages theory rests on alleged injuries that are not ultimately traceable to the breach. (ECF No. 329, pp. 53–55).

1. The Ascertainability Requirement

Because the ascertainability requirement is a threshold requirement that Plaintiffs must satisfy, this Court will address it first, before turning to the other requirements found in Federal Rule of Civil Procedure 23(a)–(b). The ascertainability requirement is an important doctrine recognized and repeatedly affirmed by the Fourth Circuit which serves to ensure that there will be an “administratively feasible [way] for the court to determine whether a particular individual” is a class member. *See Krakauer v. Dish Network, L.L.C.*, 925 F.3d 643, 658 (4th Cir. 2019) (noting that Plaintiffs do not have to “identify every class member at the time of certification”). To satisfy their proposed class definitions, Plaintiffs must demonstrate that there is an administratively feasible manner of determining whether (1) a putative plaintiff’s data was stored in the database of a customer identified in Exhibit A to Defendant’s Revised Fact Sheet from February 7, 2020 to May 20, 2020;⁷ (2) which data elements belonging to that putative plaintiff were stored in any of the 90,000 customer databases implicated in this matter; and (3) whether that data was

⁷ The parties estimate that as many as 1.5 billion constituents may have been affected by the breach. (ECF No. 317-2, p. 75). Also, the substance of Defendant’s Revised Fact Sheet referred to here as well as how it was created is explained in greater detail in Section III.A.4. In brief, Defendant’s Revised Fact Sheet contains data belonging to the named plaintiffs that Defendant confirmed was located in its customer backup files using information initially provided to it by the named plaintiffs themselves.

unencrypted. (ECF No. 293, p. 11). Further, for the proposed state law sub-classes, Plaintiffs must show: (1) the state of residency of each putative plaintiff at the time of the breach; and (2) whether specific kinds of data were impacted as required by the CMIA and CCPA.⁸ (ECF No. 329, pp. 25–28; ECF No. 329, pp. 11–12).

Plaintiffs make four arguments in support of their contention that the classes and sub-classes in this matter can be readily ascertained. They argue that each of the following demonstrates that a class is ascertainable: (1) Defendant’s ability to give notice to its customers of the breach; (2) Defendant’s ability to create its Defendant Fact Sheet which contained information about the named plaintiffs; (3) Defendant’s use of a program called Wirewheel to respond to CCPA requests; and (4) the method proposed by Curtin of identifying putative plaintiffs and identifying which of their data elements were exposed in the breach. (ECF No. 293, pp. 27–30). The Court will address each of these arguments, beginning with Plaintiffs’ argument that Curtin’s method is a reliable and helpful method of ascertaining a class. To evaluate Plaintiffs’ arguments with respect to Curtin’s method, this Court will address Defendant’s Motion to Exclude the Testimony and Report of C. Matthew Curtin as to Ascertainability (ECF No. 341) and Plaintiffs’ Motion to Exclude the Report and Testimony of Sonya Kwon (ECF No. 419).

⁸ The CMIA requires that “a patient’s medical history, mental or physical condition, or treatment” have been “viewed,” and the information in question “must be substantive.” Cal. Civ. Code §§ 56.05(i), 56.10, 56.36(b); *see also Wilson v. Rater8, LLC*, No. 20-CV-1515-DMS-LL, 2021 WL 4865930 (S.D. Cal. Oct. 18, 2021); *Sutter Health v. Superior Court*, 227 Ca. App. 4th 1546, 1555, 1557 (2014) (stating that “the records . . . [must have been] viewed by an unauthorized person”). The CCPA requires that an individual’s “first name or first initial” and “last name” together with their Social Security number, driver’s license number, credit or debit card account number in combination with the required security code or password, medical or health information, or “username or email address in combination with a password or security question and answer . . .” have been both “accessed” and “exfiltrated” by one who is “unauthorized” to do so. *See* Cal. Civ. Code §§ 1798.81.5(d)(1)(A)-(B), 1798.150(a); (ECF No. 329, pp. 17, 27–28). Further, Plaintiffs’ proposed N.Y. GBL and FDUTPA class definitions require Plaintiffs to prove that a putative plaintiff “viewed or [was] exposed to” the alleged material misrepresentations at issue. (ECF No. 293, p. 12).

2. Defendant's Motion to Exclude Curtin's Report and Testimony

Defendant has moved to exclude Curtin's report and testimony as to ascertainability, pursuant to Federal Rule of Evidence 702 and *Daubert*.⁹ (ECF No. 341). The motion has been fully briefed and is ripe for review. Further, this motion was argued on March 6, 2024 at the *Daubert* and class certification hearing held in Charleston. During that hearing, the Court heard not only from Plaintiffs' counsel, but also from Curtin himself. Further, in anticipation of the highly technical nature of both the merits of the action and the specific question of ascertainability, this Court retained a consulting expert. (ECF Nos. 433, 458). The retained expert provided this Court with valuable insight due to his experience in data science broadly and ascertainability-related matters in particular. The Court found this consultation vital to aid it in understanding the technical subject matter of Curtin's report, Kwon's report, and the issues confounding class certification as a whole.

Defendant's Motion to Exclude Curtin's report and testimony on ascertainability is granted for a number of reasons. Those reasons include Curtin's inability to provide this Court with an error rate and a statement about its occurrence consistent with generally accepted statistical practices for the Court to evaluate, his failure to sufficiently test his method, the non-replicability of his method, and his failure to sufficiently document his method so that it could be tested by Defendant's rebuttal expert. *See Daubert*, 509 U.S. at 592–95. Thus, Curtin's method cannot serve as support for Plaintiffs' contention that the proposed classes and sub-classes can be ascertained.

⁹ Curtin offers opinions both on information security and on ascertainability, which have been briefed separately by the parties. This portion of the Court's order only addresses Curtin's ascertainability opinions. Additionally, the "CISSP" certification Curtin has is "an information security certification granted by the International Information Security Consortium." (ECF No. 317-2, p. 3).

i. Curtin's Proposed Method

Plaintiffs, in their Response in Opposition to Defendant's motion to exclude Curtin's opinion and testimony (ECF No. 380), assert that Curtin's method "consists of four steps: (i) restoring customer database files (the "Queryable Databases"); (ii) generating a database called the 'Referential Index' to assist searching through the Queryable Databases; (iii) obtaining basic information from class members; and (iv) executing a search through the Queryable Databases for data elements relating to the class members." (ECF No. 380, pp. 9–10). It is worth noting that this presentation of Curtin's method differs from Plaintiffs' description of Curtin's method in their Motion to Certify a Class. (ECF No. 329). In Plaintiffs' Motion to Certify, they state that Curtin can "identify class members by querying the email addresses present in the data set provided by Blackbaud to Plaintiffs' counsel during discovery."¹⁰ (ECF No. 293, p. 27). They also state that Curtin can "query the affected data elements for any class member, which will permit Plaintiffs to identify which Classes and Sub-classes a class member belongs to." (ECF No. 293, p. 27).

Notably, Plaintiffs appear to abandon their initial argument that Curtin can identify potential class members using the email addresses present in the customer backup files, which underscores the moving-target nature of Plaintiffs' arguments regarding Curtin's method. Accordingly, Plaintiffs have made determining what steps Curtin's method entails no easy feat. Plaintiffs' description of Curtin's method has changed significantly from the time the Motion to

¹⁰ Curtin stated at his deposition that he believed that a class member could be identified, using his method, simply via the presence of their email address within a Referential Index consisting of data compiled from the restored customer backup files that were accessed during the data breach. (ECF No. 387-31, pp. 5-6). Curtin was specifically asked: "[I]t's your testimony that the email addresses . . . that you found that are in your referential index, those would be the members of the class?" (ECF No. 387-31, p. 6). Curtin responded by saying: "I was given the information from Blackbaud that was identified as part of what was exposed in the data breach, so who we have in there should be class members, yes." (ECF No. 387-31, p. 6). He also indicated that additional information from class members would only be needed to "answer[] specific questions . . . [such as] whether someone is a resident of California, for example . . . but we've already identified that person as part of the class." (ECF No. 387-31, p. 5). Plaintiffs continually insist that this is not what Curtin's method will do, but Defendant's arguments on this point were made in direct response to Curtin's own testimony.

Certify and Curtin's accompanying report were first filed on December 16, 2022 and February 7, 2023 respectively to the *Daubert* and class certification hearings held in March of 2024. (ECF Nos. 293, 317-2). Plaintiffs initially asserted that Curtin could "identify class members by querying the email addresses present in the data set" Defendant provided to Plaintiffs. (ECF No. 293, p. 297). Such an assertion is susceptible to only one interpretation. It is clear that, at one point, Plaintiffs believed that Curtin could identify class members themselves using the email addresses present in the customer backup files. Plaintiffs no longer make this argument in their Response in Opposition to Defendant's Motion to Exclude Curtin's Report and Testimony. (ECF No. 342). Instead, Plaintiffs state that Curtin's Referential Index will **not** be used "to generate a list of email addresses to uniquely identify Class and Subclass members."¹¹ (ECF No. 380, p. 11). Plaintiffs now assert that Curtin will "use[] information provided by a putative class member . . . to search for that person, identify her data that was exposed . . . and then confirm that she is a member of the class." (ECF No. 380, p. 7). In light of this change, this Court will evaluate Curtin's method as described in Plaintiffs' Response in Opposition. (ECF No. 380).

Importantly, Plaintiffs have stated both at an earlier Case Management Conference and at the *Daubert* and class certification hearings held in March of 2024 that the report and testimony provided by Curtin are unnecessary to this Court's determination of whether a class is ascertainable in this case. (ECF No. 432; ECF No. 496, p. 20 (stating that "Curtin's software code is icing on the cake" and is "ancillary" to Plaintiffs' other proposed methods of ascertaining a class)). That being the case, this Court's ruling on Defendant's Motion to Exclude (ECF No. 380) should do

¹¹ Plaintiffs suggest that Defendant misunderstood Curtin's report and their initial Motion to Certify a Class, which is not the case. Plaintiffs plainly stated in their Motion to Certify that Curtin was able to "identify" class members using the email addresses present in the customer backup files and organized within the Referential Index. (ECF No. 293, p. 27). The Plaintiffs' decision to shift course does not render Defendant's initial Opposition a misunderstanding. (ECF No. 329).

little, if anything, to prejudice Plaintiffs in light of their clearly-articulated position that Curtin’s opinions on ascertainability are unnecessary to this Court’s class certification analysis. (ECF No. 432, pp. 11–14). However, this Court is doubtful of Plaintiffs’ suggestion that expert testimony is unnecessary for an ascertainability determination in a case as complex as this one.

ii. Curtin’s Proposed Method is Not Reliable

First and foremost, Curtin’s method is unreliable and unhelpful to this Court in light of Curtin’s failure to provide an error rate for this Court to evaluate. *See Daubert*, 509 U.S. at 594. The unreliability of Curtin’s method is further underscored by the fact that his method (specifically his Referential Index) cannot be replicated, several of the steps he has proposed have not been sufficiently tested, Curtin has failed to identify the “final product” of step two of his method, and Curtin has not demonstrated that his method can be scaled to operate accurately across classes and sub-classes consisting of as many as 1.5 billion putative plaintiffs. *See Daubert*, 509 U.S. at 593–95.

(a) Curtin’s Failure to Properly Test or Provide an Error Rate

Two of the *Daubert* factors are the “known or potential rate of error” and “whether a theory or technique . . . can be (and has been) tested.” *Daubert*, 509 U.S. at 593–94; *see also Nease v. Ford Motor Company*, 848 F.3d 219, 232 (4th Cir. 2017) (noting that an expert’s opinion, although “plausible,” was rendered unreliable and “no more than a hypothesis” due to the expert’s failure to “validate it with testing” and to thereby provide the court with a “potential error rate”). Courts may also “conclude that there is simply too great an analytical gap between the data and the opinion proffered.” *Gen. Elec. Co.*, 522 U.S. at 146. Further, and particularly relevant here, “[n]othing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert.” *Id.*

In this instance, Curtin has provided this Court with mere *ipse dixit* upon which it cannot rely by giving an opinion that his method can be effectively used on a large scale and across a wide number of data sources without providing this Court any way to evaluate the accuracy of the results his methods will produce. (ECF No. 317-2, p. 142 (stating that Curtin’s “automated inspection process can be applied to any number of files . . .”). This Court can certainly envision a world in which Curtin’s method as proposed could produce the results he is describing. However, this Court is unwilling to accept Curtin’s word that his method will operate effectively and accurately without scientific testing that supports Curtin’s assertions that his method is accurate and automatable across the millions of putative plaintiffs that this class would potentially involve.¹² To date, Curtin has tested his method on three named plaintiffs and has fully documented the queries he used for only two of those individuals.¹³ (ECF No. 317-2, pp. 62–63). To insist that a method will work across eleven customizable products, approximately 13,000 customers, approximately 90,000 customer backup files, and widely varied data belonging to tens of millions of putative plaintiffs because it worked to identify information for three named plaintiffs across two products is a bridge too far. This Court is not required to accept Curtin’s *ipse dixit* that his method will work to ascertain such a large and potentially complex set of classes and sub-classes without any sort of error rate or testing to support that assertion. *See Daubert*, 509 U.S. at 594.

There are numerous other components of Curtin’s method that he has not tested and for which he has not otherwise provided any indicator of accuracy that require this Court to conclude

¹² This Court asked Curtin at the *Daubert* hearing on Defendant’s motion to exclude his report and testimony on ascertainability whether he was aware of the rate at which data elements might be accidentally “mixed and matched” or improperly conflated across putative plaintiffs. (ECF No. 494, p. 98). Curtin acknowledged that such conflation was possible and stated that he was not aware of the rate at which it might occur, suggesting that his job would merely be to “look . . . up” data as it is presented to him. (ECF No. 494, p. 98).

¹³ The word “queries” refers to code that is written to perform a particular search. (ECF No. 317-2, pp. 74–75 (stating that Curtin, for instance, is “able to query” certain email addresses and referring to the tools he used to identify email addresses as “queries”)).

that his method is unreliable under Rule 702 and *Daubert*. First, Curtin has not written queries for each type of data element that would need to be identified, such as medical records, credit card records, drivers' license numbers, and license tag numbers. *See* (ECF No. 342, pp. 30–31 (noting that Curtin only identified a name, email, phone number, and encrypted social security number in the searches he ran)). He simply asserts that he is able to do so, using the queries he did write as support for this proposition. (ECF No. 330-3, p. 55). Most glaring of these absences is the absence of thorough attempts to write or test queries for PHI. (ECF No. 342, pp. 30–31).

Second, Curtin ran no tests, or at least provided no record of any testing, to demonstrate his ability to accurately identify the state of residency and the necessary categories of data belonging to putative plaintiffs in the proposed state law subclasses. He merely stated that he has identified individuals who reside in California, New York, and Florida but has not validated those findings. (ECF No. 329-3, p. 68–71; ECF No. 317-2 p. 74 (stating merely that Curtin can “confirm an individual’s state of residence in the data if that information is provided”)).¹⁴ Importantly, Defendant has already demonstrated that two named plaintiffs’ addresses within the customer backup files do not match their address at the time of the breach. (ECF No. 329, p. 25 (noting that one named plaintiff lived in New York in 2020 but had a New Jersey address in the backup files and that another named plaintiff lived in Colorado in 2020 but had a Mississippi address in the backup files)). Residency is just one requirement of several that the four state laws in question impose for a plaintiff to state a claim. The CMIA and CCPA also require that specific kinds of

¹⁴ Defendant correctly notes that Curtin did not look beyond surface-level indicators of residency found in some of the data he inspected. Curtin also did not account for the fact that constituents may have moved after data indicating their residence was initially stored by a Blackbaud customer, and not all customers necessarily update their data regularly. (ECF No. 329, p. 25). Thus, the mere presence of a street address, state of residence, or other piece of information stored in a customer backup file without more does not show that an individual satisfies the state law subclass definitions proposed by Plaintiffs. Curtin has provided no method through which a putative plaintiffs’ state of residence at the time of the breach can be validated. Additionally, “the resolution of [the] legal residency requirement would require . . . individualized analysis” that this Court does not believe can reasonably be undertaken for reasons that will be further discussed below. (ECF No. 329, p. 25).

data be exposed for an individual to be eligible for statutory damages. (ECF No. 329-3, pp. 70–72). Curtin did not search for or identify any “medical information” as defined by the CMIA (ECF No. 329, pp. 25-27) or “personal information” as defined by the CCPA, nor did he write any queries or otherwise demonstrate how he would do so. (ECF No. 329, pp. 27-28). Again, this will not do—Curtin’s mere word that he is able to identify sub-class members’ state of residence and specific kinds of data depending on the requirements of each state law at issue does not suffice under Rule 702 or *Daubert*.

Third, Curtin says he can use “the information provided by [a] class member” to identify their affected data elements, but he only did so using first and last name and email addresses in his report. (ECF No. 317-2, p. 141). Curtin relied on information from the Plaintiff Fact Sheets to run his searches, which he will not necessarily have were he to apply his method on a larger scale. He has not demonstrated how his method would work using other data elements even once, let alone at a large scale. Further, even the elements he did search for often returned a large number of results that he has not demonstrated an ability to filter to eliminate incorrect or inaccurate information that is not actually related to a class member. (ECF No. 494, pp. 46–48 (explaining that Curtin’s Referential Index returned 179 street addresses in response to a query for one email address, and that another query for an email address returned over 1600 unique street addresses in addition to multiple different names)). This Court has more fully addressed the problem of what Curtin will use to search for putative plaintiffs’ data below. *See infra* Section III.A.4. Regardless, the undeveloped nature of this step as set forth in Curtin’s method is mere *ipse dixit* that this Court will not consider.

Fourth, Curtin has not inspected or tested how his queries would work across all or even a majority of Defendant’s products, choosing to test it on just two. (ECF No. 342, pp. 29–30; ECF

No. 494, p. 92 (stating that Curtin “did not make choices” and “did not select particular things . . . [he] made no choice based on [a particular] product”)). Defendant’s products are customizable, such that customers who are using the same product can store the same kinds of data in different encrypted and unencrypted fields. (ECF No. 342, pp. 32–34). As Defendant points out, Raiser’s Edge NXT was used by two different Blackbaud customers to store constituent data in notably different ways. Stetson University stored one former named plaintiff’s social security number in an “unencrypted free form field” not meant for the storage of social security numbers, while Mercy Health Services stored another named plaintiffs’ social security number in an encrypted field. *Id.* at 33.

Another illustrative example of the potential for variety in the data elements stored by one single customer using one product can be found in the fact that Trinity Health stored treating physician names and dates for one named plaintiff and stored insurer and donation history information for another.¹⁵ (ECF No. 324, pp. 33–34). This level of variation in method of data storage and types of data stored for just one individual or using just one product suggests that an even greater degree of variation exists across all products and all customer databases. Curtin admittedly does not know how varied customer’s uses of Defendant’s products are, nor how the customizations implemented by Defendant’s customers will impact his ability to query the customer backup files, since he only restored the most recent backup file for the customers he dealt with. (ECF No. 494, pp. 90, 92). Given that customers can and have stored data differently within just one product or for just one person, the amount of variation that is possible across all 11 products and all 90,000 customer backup files is not insignificant. (ECF No. 329-3, pp. 56–57).

¹⁵ The opposite was not true, meaning that Trinity Health did not store treating physician names and dates for the individual for whom they stored insurer and donation history information, and they did not store insurer and donation history information for the individual for whom they stored treating physician names and dates. (ECF No. 324, pp. 33–34).

Thus, Curtin has not conducted sufficient testing to support his assertion that those differences in customization of a single product by different customers and differences in kinds of data stored will not affect his method's utility.¹⁶

(b) Curtin's Method Cannot Be Replicated

The Fourth Circuit, along with many other circuits, has acknowledged that replicability is a strong indicator of reliability and that the inability to replicate an expert's methods may indicate that the expert's methods are not reliable. *See, e.g., Ruffin v. Shaw Indus., Inc.*, 149 F.3d 294, 297–99 (4th Cir. 1998). Curtin's method of ascertaining class members relies heavily upon the use of a Referential Index, which is a database that Curtin built by restoring and compiling the sample of Blackbaud customers' data Defendant provided to make that data easier to search for specific data elements. *See* (ECF No. 317-2, pp. 133–34). Plaintiffs state that Curtin's Referential Index works “by storing, for a given piece of data about a Class member in the Referential Index—for example an email address—the identity of which Queryable Databases that email address appears in, where in those [databases] additional data elements associated with the email address can be found . . . and the specific search queries that Curtin's software executed to locate those data elements.” (ECF No. 410, p. 10).

The construction of a Referential Index using the restored customer backup files is step two in Curtin's method. (ECF No. 380, pp. 10–11). Importantly, while developing his method and drafting his report, Curtin created three separate Referential Indexes. (ECF No. 329-3, pp. 41–42).

¹⁶ Defendant points out that Plaintiffs rely almost exclusively on Curtin's deposition testimony to support their arguments that Curtin's method will work despite differences in customers' use of a single product and despite differences across Defendant's products in general. Specifically, Defendant notes that Curtin “did not develop and produce a script that could be used to search for even one putative class member” in the 20 Raiser's Edge backup files that Curtin had access to, “much less a script that will work to identify data elements associated with all putative class members whose data was stored in a Raiser's Edge backup file.” (ECF No. 400, pp. 12–13). Deposition testimony without anything more to support it is the precise sort of *ipse dixit* that courts are not permitted to rely upon when deciding whether to include or exclude an expert's opinion.

His report does not state which Referential Index should be relied upon by Defendant when testing his method, nor was he able to firmly point to which Referential Index should be relied upon when asked to do so at his deposition. *See* (ECF No. 330-3, pp. 45–46). In fact, Curtin contradicted himself in the two different sessions of his deposition, unable to identify whether Instance 1 or Instance 3 of his Referential Index was the version Defendant should be using to attempt to replicate his results. *See* (ECF No. 330-3, p. 47; ECF No. 387-31, pp. 8–10).

Further, Curtin failed to provide clear instructions to allow the Defendant’s expert, Sonya Kwon (“Kwon”), to recreate his Referential Index for testing purposes. (ECF No. 329-3, pp. 39–40; ECF No. 342, pp. 24–25). Plaintiffs now attempt to restyle Curtin’s Referential Indexes as merely an “intermediate component” of Curtin’s method rather than the final product and insist that all Kwon needs to test Curtin’s method is Curtin’s scripts.¹⁷ (ECF No. 380, pp. 11–12). This will not do. Kwon should have been afforded the opportunity to inspect the “final product” of Curtin’s scripts so that she could determine whether the scripts or underlying data inputs were edited in response to the final product and whether Curtin applied any interpretations or assumptions to the final product. Plaintiffs’ insistence that there is no real “final” Referential Index and that no final product aside from scripts is needed to assess this important part of Curtin’s

¹⁷ Plaintiffs explain that “SQL (‘Structured Query Language’) is a programming language used to manage data held in relational databases” and that “SQL scripts” or “scripts” are “pieces of code written to run in SQL.” (ECF No. 410, p. 10). Notably, the scripts produced by Curtin were written after his report was initially submitted. Curtin was unable to provide a clear explanation at the March 6-8 hearings for why the scripts he produced in discovery that he ostensibly used to produce his Referential Indexes post-date the creation of those Referential Indexes. Specifically, this Court asked Curtin: “Why were the three Python scripts that you said you used to create your referential indexes written after your report was filed?” (ECF No. 494, p. 95). Curtin replied: “The process of showing how that was going to be easily reconstructed had to do with the process of handing the environment over to the other side. I was expecting that we would have a production process that would go along with that. And so, to ease that process of production, we wanted to put that together for opposing counsel.” (ECF No. 494, pp. 95–96). Far from a model of clarity, Curtin’s response leaves the Court with more questions than answers on this point. It certainly fails to sufficiently explain why the scripts that were purportedly used to create the Referential Index(es) post-date the submission of Curtin’s report. Accordingly, this Court is convinced that Kwon was hampered in her ability to test and critique Curtin’s method, since she lacked his final product and was not given scripts that pre-date the submission of his expert report to use in her attempts to replicate his method.

method is unavailing.¹⁸ The lack of a “final product” significantly hampered Kwon’s ability to evaluate and test Curtin’s method and has thereby prevented this Court from knowing how Curtin’s opinions about this key step in his method relate to the output of those scripts. Stated differently, Kwon was handicapped from the outset because Curtin gave her a “proof of concept” without clear instructions as to implementation or utilization and without clear conclusions or output against which to check either his work or hers. However, regardless of whether Curtin’s “final product” is available to Kwon, the Referential Indexes she did inspect are undoubtedly similar to Curtin’s final product. Kwon’s criticisms of those Referential Indexes are thus helpful to this Court in its evaluation of Curtin’s methods. Ultimately, Curtin’s failure to identify a final Referential Index and his failure to provide scripts that pre-date the production of his report significantly hampered Kwon’s ability to test and replicate his methods, as she was missing important pieces of his work.

Curtin’s non-production of a final Referential Index prevented Kwon from identifying the exceptions or anomalies that potentially arose when Curtin was using the scripts that he wrote, seeing how Curtin resolved any such issues and altered his code to address them, or knowing the amount of manual work that Curtin and his team had to expend across the three named plaintiffs. Muddled documentation and code writing alone may not impact the accuracy of Curtin’s methodology, but his results cannot be fully tested and validated as a consequence of the confusion around the relevant “scripts” and the application of his methodology. Ultimately, these failures on

¹⁸ This Court finds it hard to believe that an expert retained primarily to assist in proving that a class can be ascertained would fail to create a final version of the Referential Index as he workshopped his method and that he ultimately or primarily relied upon when forming his opinions—whether that be Instance 1, Instance 3, or some other version of his Referential Index altogether. (ECF No. 494, p. 95 (stating that Curtin never “buil[t] an ultimate referential index”). Curtin maintained an inconsistent position on this point even at his deposition, insisting on the one hand that he had not created a final Referential Index, but also stating that his work was “done . . . as to feasibility.” (ECF No. 411-1, p. 9; ECF No. 330-3, p. 50). The lack of clarity from Plaintiffs and Curtin on this point is unhelpful to this Court and further underscores the overall lack of reliability in Curtin’s method.

Curtin's part are just one of several reasons that his method is unreliable and fails under Rule 702 and *Daubert*.

(c) Curtin Has Not Demonstrated That His Method Can Be Automated

The Fourth Circuit's ascertainability requirement, first recognized in *Hammond v. Powell*, provides that the members of a class must be "readily identifiable." 462 F.2d 1053, 1055 (4th Cir. 1972). Curtin suggests that his method can be effectively and correctly implemented for tens of millions of putative plaintiffs across thousands of disparately constructed customer databases while he has, again, only tested his method on three individuals and two products. (ECF No. 317-2, p. 141; ECF No. 494, pp. 92–93). This Court might be inclined to accept Curtin's representations regarding his method's automatability if he had provided it with an acceptable error rate to evaluate. However, the absence of any assurance that Curtin's methods are reasonably accurate renders this Court unable to find that Curtin's methods can work well at the large scale for which they would need to be deployed if the proposed classes and sub-classes were certified. This Court is not merely concerned with whether Curtin's scripts can be modified to operate at a larger scale; it is concerned with whether Curtin's scripts will operate at a large scale accurately and whether a high degree of manual intervention will be necessary to that end. In the absence of any testing or scientific error rate, this Court is unable to simply take Curtin's word for it that his method can work at the necessary scale.¹⁹ Accordingly, Curtin cannot demonstrate that his method is "administratively feasible" and that it can be effectuated without "extensive and individualized fact finding." *Spotswood*, 2019 WL 498822, at *6.

¹⁹ Further, the fact that the two sets of scripts that Curtin produced for Plaintiffs Eisen and Kamm were only effective when used to search for data belonging to the individual for whom they were originally written suggests that Curtin's method cannot be scaled as easily as he suggests. See (ECF No. 342, pp. 14–15). If the two sets of scripts Curtin produced required manual intervention and tailoring just to locate data for two individuals, this Court has no reason to believe that Curtin's method could be used to locate the data of millions of plaintiffs without a significant and ultimately non-feasible level of human intervention being necessary. Curtin himself has acknowledged that his method would take at least "several thousand hours" to implement across the entire class. (ECF No. 494, p. 100).

(d) Curtin Has Not Explained How He Would Validate His Results

Implicit in the nature of Curtin's method and in the nature of this case is the need to determine whether the data profiles Curtin's method generates are an accurate compilation of a particular plaintiff's exposed data elements.²⁰ Curtin tested his method on three named plaintiffs, but he did not indicate whether or how he verified the accuracy of the identified data elements in his test searches for those three individuals.²¹ Notably, Curtin had the benefit of Plaintiff Fact Sheets containing information provided by the named plaintiffs that allowed him to begin his searches using email addresses that he knew belonged to the individuals in question along with the kinds of specific data elements that they believed had been exposed, a tool he would not have if he were looking for data belonging to non-named plaintiffs.²² These fact sheets enabled him to be sure the email address he was using belonged to the correct individual, as opposed to an organization or another person altogether, which Curtin has acknowledged is a possible problem when using his method. (ECF No. 494, p. 98). The fact sheets also enabled Curtin to know what kinds of data to look for and gave him the ability to confirm that certain data elements were correct, since he had the named plaintiffs' self-verified information to compare his method's results to. *See, e.g.*, (ECF No. 387-15, pp. 4–8 (listing named plaintiffs' email address, date of birth, street

²⁰ Although Plaintiffs are not required to actually identify who all of the class members are at this stage of the litigation, they still must demonstrate **how** the nationwide classes and sub-classes would be ascertained. This requires Plaintiffs to show how they will verify that the data profiles Curtin's method will produce (1) belong to the person they are purported to belong to and (2) contain a correct list of data elements. Otherwise, this Court has no way of being sure that Curtin's method is not (1) misidentifying individuals by using an email address or other identifier that does not actually belong to a particular plaintiff or (2) including data elements that do not belong to a particular plaintiff in that individual's profile or leaving out data elements that should be included in a particular individual's profile. Incorrect data profiles are not helpful to this Court and would not satisfy the Plaintiffs' proposed class definitions.

²¹ Curtin has stated that determining whether the data within the Referential Index itself is accurate is "beyond the scope of what [he was] engaged to do." (ECF No. 342, p. 26). Further, he has stated that validating the results of his method "would be part of the claims processing." (ECF No. 494, p. 95). Ultimately, he does not appear to have contemplated the need for data validation, especially of the data profiles his method would produce, in the course of developing his method.

²² An explanation of how and why the Plaintiff Fact Sheets were created can be found in Section III.A.4.

address, and the Blackbaud customers to whom the individual believed they had given their data, along with an estimate of which kinds of information were given to each customer)).

Ultimately, Curtin has stated that validating the data profiles his method produced was beyond the scope of what he was asked to do. (ECF No. 342, p. 26). This leaves the Court with no way of knowing how Curtin would confirm that (1) a given set of data elements all belong to the person associated with the email address or other information that is used as the “key” or starting point in his searches; that (2) no data elements belonging to that person and exposed in the breach were left out if he were to run his searches on a larger scale; and more importantly that (3) no data elements belonging to someone else have been included in the set of data elements Curtin’s method produces. Curtin himself acknowledges that conflation of data across putative plaintiffs is possible and states that he has no idea at what rate that might occur. (ECF No. 494, pp. 98–99). In other words, in the absence of confirming data that can be systematically used to ensure the accuracy of search results, Curtin’s results may be subject to manual intervention to validate its output, which is not viable on a broad scale. As a result, when Curtin’s method is applied more broadly, he might conflate data elements associated with multiple different people, improperly attribute data elements to a person, or leave out data elements that were exposed in the breach. Because Curtin has not otherwise attempted to provide this Court with an error rate to evaluate, this Court is unable to rely on Curtin’s assumption that the results his method would produce would be accurate without a proposed method of validation when used at the scale that this case would require.²³

²³ An issue that especially troubles this Court is the fact that Curtin does not address at all the risk that PII or PHI might be mismatched and ultimately transmitted to the wrong putative plaintiff while implementing his method. Because Curtin has not validated the data himself or demonstrated that it can be validated without involvement from the putative plaintiffs themselves, there remains a distinct possibility that PII or PHI belonging to one person could be inadvertently transmitted to another person because of Curtin’s method.

(e) Curtin Was Not Limited by the 100 Customer Sample

Curtin and Plaintiffs have complained that Curtin was unable to effectively test his method because he was limited to a sample of 100 customers' backup files. (ECF No. 400, p. 15). This Court's holding that Curtin's methods are unreliable is unrelated to the limited sample of customers that Curtin was using. Rather, as Defendant points out, the issue is that Curtin failed to properly test his method or provide an error rate with respect to the "backup files he did have." (ECF No. 400, p. 25). Curtin could have conducted additional testing and statistical analysis of the data present in the 100-customer sample to thoroughly examine whether his proposed method is producing the kind of output that he says it will and to search for unknown risk of error. Curtin chose not to conduct any testing beyond running scripts for three named plaintiffs or to otherwise search for and mitigate risk of error, overinclusion, underinclusion, and conflation of individuals' data. That decision has rendered his method unreliable regardless of the size of the sample he was using.

Thus, Defendant's Motion to Exclude the Report and Testimony of C. Matthew Curtin is granted because his proposed method of ascertaining a class is unreliable and unhelpful to this Court in light of his failure to satisfy numerous *Daubert* factors. Several aspects of Curtin's method were not properly tested, he failed to provide this Court with a "potential rate of error" to evaluate, he failed to consider alternative explanations for his conclusions, and he extrapolated from numerous accepted premises to unfounded conclusions. *See Daubert*, 509 U.S. at 593–94; *see also Daubert*, 43 F.3d at 1317. This Court does not take lightly its responsibility to serve as a gatekeeper and to decline to rely upon unreliable expert testimony or to allow such to be presented to a jury. Curtin's report falls short of the standards the Fourth Circuit and the Supreme Court have set for expert's reports and testimony, and it should be excluded. Defendant has separately sought to

exclude Curtin's report and testimony on the grounds that he is not qualified to offer ascertainability opinions and that he impermissibly relies on the work of others. Because this Court finds Curtin's method to be unreliable under *Daubert* and Rule 702, Defendant's motion to exclude the portions of his report and testimony that address his method of ascertaining a class is granted on that ground alone. Because Curtin's proposed method is excluded under *Daubert*, it cannot serve as support for Plaintiffs' argument that the proposed classes and sub-classes are ascertainable.

3. Plaintiffs' Motion to Exclude the Report and Testimony of Sonya Kwon

In ruling on Defendant's motion to exclude Curtin's report and testimony as to ascertainability (ECF No. 341), this Court must also rule on Plaintiffs' Motion to Exclude the Report and Testimony of Sonya Kwon. (ECF No. 419). The motion has been fully briefed and is ripe for review. Further, this motion was argued on March 6, 2024 at the *Daubert* and class certification hearing held in Charleston. Sonya Kwon is Defendant's rebuttal expert who, in pertinent part, offers opinions seeking to counter Curtin's expert report and testimony on ascertainability. Plaintiffs do not seek to fully exclude Kwon's report and testimony. Instead, they specifically seek to exclude two sets of opinions that Kwon offers on (1) the replicability of Curtin's method and (2) the usefulness of William Wecker's sample in "extrapolat[ing] conclusions relating to the size or properties of the class and subclasses . . .". (ECF No. 410, p. 8). This Court's ruling is limited to Plaintiffs' motion to exclude Kwon's opinion that step two of Curtin's method, his creation of a Referential Index, is not replicable. Plaintiffs contend that this portion of Kwon's report and testimony is irrelevant and should be excluded because "[Kwon] compared Curtin's final implementation of his method to previous, prototypical versions" of his method, "which says nothing about whether his method is replicable." (ECF No. 410, p. 8).

Plaintiff's motion to exclude Kwon's opinion on the replicability of Curtin's Referential Index is denied. Her opinion that she was not able to create a Referential Index, using the instructions and scripts that Curtin provided, that matched any version of a Referential Index that Curtin created suggests that his method is non-replicable and ultimately unreliable. Thus, her opinion is both relevant and helpful and should not be excluded.

Plaintiffs argue that Kwon's opinion—that step two of Curtin's method, his creation of a Referential Index, is not replicable—is “fundamentally flawed and irrelevant” because it is premised on a misunderstanding of Curtin's method. (ECF No. 410, p. 14). Specifically, Plaintiffs say that Kwon is errantly comparing “legacy, prototype versions” of Curtin's Referential Index with the “results” of Curtin's method, when she should simply be “checking whether Curtin's method reliably generates the same results every time.” (ECF No. 410, p. 14). Defendant counters that Kwon's opinion “directly rebuts” Curtin's own words found in his report and is therefore highly relevant. (ECF No. 448, pp. 8–9). Defendant points out that Curtin's report clearly states that he created “an index of data,” which he refers to as “**the** Referential Index.” (ECF No. 448, p. 9) (emphasis added). Curtin also states that he “is able to store the specific SQL queries used to build the Referential Index, which would allow for someone else to re-create the Referential Index using my same methodology.” (ECF No. 448, p. 9). Thus, Defendant argues that Kwon's rebuttal opinion that Curtin's Referential Index “cannot even be identified, much less replicated” is helpful and relevant to this Court's evaluation of Curtin's opinions and testimony. (ECF No. 448, p. 9).

Plaintiffs' argument that Kwon's replicability opinion is based on a flawed premise and is therefore unhelpful to this Court under *Daubert* and Rule 702 falls flat. As an initial matter, Curtin's opinion repeatedly refers to “a” Referential Index or “the” Referential Index, and the use of this Referential Index is important to his method in at least two ways—first as a method of

identifying putative plaintiffs (an argument that Plaintiffs appear to have abandoned) (ECF No. 410, pp. 9–10 (stating that “[t]he basic purpose of the Referential Index is to make Curtin’s search method work more quickly and efficiently . . .”)), and second as a method of finding the data elements belonging to putative plaintiffs so that they can be compiled into a profile for each plaintiff. (ECF No. 410, pp. 10–11). Some of Kwon’s criticisms of Curtin’s Referential Index are premised on the now-abandoned opinion offered by Curtin that his Referential Index can be used to “identify[] the constituents whose information is stored in the backup files” using their email addresses. (ECF No. 317-2, pp. 2, 141; ECF No. 387-31, pp. 4–5 (stating that Curtin’s process involves “go[ing] through and identify[ing] email addresses . . . [and thereby] identify[ing] that person as part of the class”)). This Court will not exclude opinions offered by Kwon in response to an initially raised and now-abandoned argument simply because Plaintiffs now choose to make different arguments about the Referential Index. Further, Kwon’s opinions regarding the replicability of his Referential Index are still relevant and helpful and should not be excluded for the reasons addressed below. *See Sardis v. Overhead Door Corp.*, 10 F.4th 268, 282 (4th Cir. 2021) (stating that a “court must satisfy itself that the proffered testimony is relevant to the issue at hand” when “a party challenges an opposing expert’s testimony as irrelevant”).

Despite Plaintiffs’ decision to abandon Curtin’s opinion that his Referential Index enables him to identify members of the class using their email addresses, the functionality and replicability of the Referential Index are still important to this Court’s assessment of Curtin’s overall method, as his Referential Index remains a key component of that method. (ECF No. 410, pp. 9–10). Even under Plaintiffs’ revised description of Curtin’s method, his Referential Index purportedly “gives [Curtin] information about what data elements are associated with the piece of data he queries, plus a ‘map’ for where to find additional data elements relating to the email address.” (ECF No.

410, p. 10). Thus, whether Curtin’s Referential Index is testable and replicable under *Daubert* are relevant and appropriate points for Kwon to address in her position as a rebuttal expert. *See Ruffin*, 149 F.3d at 297 (stating that whether a “technique . . . has been tested and independently validated or replicated” is a “‘key question’ in determining whether a technique can be considered reliable scientific knowledge”). For instance, Kwon importantly pointed out at the March 6-8 hearings that the Referential Index present in Instance 1 contained an incorrect association between individuals and “encrypted credit card information.” (ECF No. 494, pp. 45–46). Curtin intended to “connect encrypted credit card information to putative class members” and instead “associated encrypted credit cards to people who only paid using checks and never provided any credit card information.” *Id.* Without a final Referential Index to inspect, Kwon cannot tell whether Curtin remedied errors of this sort, of which Kwon asserts there are “tens of thousands,” or whether they persisted.

This Court is perplexed by Plaintiffs’ assertion that draft versions of Curtin’s Referential Index are present in the virtual environment that he used, but no final version of that Referential Index is present within the virtual environment, nor was one ever created at all. (ECF No. 410, p. 12; ECF No. 387-31, pp. 8–9 (confirming that a “referential index” is in both database Instance 1 and in database Instance 3)). Curtin initially anticipated that someone else would need to re-create his Referential Index, despite Plaintiffs’ contention that Kwon’s attempts to do so were errant. (ECF No. 317-2, p. 136). However, Curtin appeared to equivocate at his deposition when asked how his Referential Index could be replicated. Defendant specifically asked Curtin at his deposition whether it was his testimony that none of the Referential Indexes present on either Instance 1 or Instance 3 could be used to replicate his results. (ECF No. 387-31, p. 8). In response, Curtin said: “You can certainly query database instance 1 and the referential index there,” while noting that certain databases may have “failed in the restoration process.” (ECF No. 387-31, pp.

8–9). In sum, Kwon’s opinion that Curtin’s Referential Index is not replicable, and that his method is therefore unreliable, is directly in response to a step in Curtin’s method that he proffered in his report and addressed at his deposition. Plaintiffs cannot insist that Curtin never finalized this step in his method and thereby immunize his attempted creation of a Referential Index from critique.

Kwon’s efforts to use the scripts provided by Curtin on the same data that he used to replicate a Referential Index that resembled one of the three that he had built in his virtual environment constitute proper rebuttal testimony. *See United States v. Stitt*, 250 F.3d 878, 897 (4th Cir. 2001) (stating that rebuttal evidence is “evidence given to explain, repel, counteract, or disprove facts given in evidence by the opposing party”) (cleaned up). Specifically, Kwon attempted to recreate Curtin’s Referential Index automatically using a set of Python scripts designed for that task and manually by running the five SQL scripts that Curtin directed her to use to see if either would produce a Referential Index that corresponded with either of the Referential Indexes that Curtin built.²⁴ (ECF No. 329-3, pp. 41–50). She observed that her attempted recreations contained significantly different numbers of rows and tables of data and different numbers of email addresses than either of Curtin’s Referential Indexes. (ECF No. 329-3, pp. 47–50). In light of that discovery, she offered an opinion that Curtin’s Referential Index is not replicable and that his method is ultimately unreliable. Plaintiffs cannot require Kwon to only test Curtin’s method in the manner that they would prefer, by running his scripts with no frame of reference to compare the product of those scripts to. Kwon’s analysis of Curtin’s Referential Indexes and her opinion that his method is flawed partly because his scripts do not produce a

²⁴ Python scripts are “code” that Curtin instructed Kwon to use to “rebuild[] the database.” (ECF No. 329-3, p. 44). The Python scripts referred to here were meant to “restore the Backup Files to create the Queryable Databases and then extract data from the Queryable Databases to create the Referential Index.” *Id.* Kwon also explained that she “used Mr. Curtin’s Python script to attempt to re-create the Referential Index.” When she “observed that Mr. Curtin’s Python script produced significantly less results than what he had recorded in the Referential Index on Instance 1,” she “ran each of the five SQL scripts manually based on the order specified in his Python script.” (ECF No. 329-3, p. 46).

uniform and consistent Referential Index when tested is both relevant and helpful. Her efforts fall squarely within the purview of a rebuttal expert, and they provide important insight into whether a key step in Curtin’s method works as described in his report and deposition testimony.²⁵

Lastly, the argument that Kwon’s ability to “run[] Curtin’s scripts . . . [on] the same set of customer databases . . . [and] produce the same Referential Index” suggests that Curtin’s own method is replicable is without merit. (ECF No. 410, p. 16). It makes perfect sense that Kwon arrived at the same results when she and her team ran the same set of scripts on the same data set multiple times. Kwon’s ability to replicate *her own* results when using Curtin’s scripts in no way indicates that Curtin’s Referential Index is similarly replicable, nor does this aid Plaintiffs in meeting their burden of proof regarding the reliability of their ascertainability expert’s opinions. Defendant correctly notes that *Daubert*’s testability factor primarily requires that “‘someone else using the same data and methods . . . be able to replicate the results.’” *See City of Pomona v. SQM N. Am. Corp.*, 750 F.3d 1036, 1047 (9th Cir. 2014) (quoting *Zenith Elecs. Corp. v. WH-TV Broad. Corp.*, 395 F.3d 416, 419 (7th Cir. 2005)); *see also Ruffin*, 149 F.3d at 297–99. The backwards suggestion that a rebuttal expert’s own internal consistency when attempting to re-create an opposing party’s expert’s results demonstrates replicability is untethered to the law and does nothing to advance Plaintiffs’ arguments in favor of their motion to exclude Kwon’s replicability opinion. Kwon’s ability to produce a Referential Index using scripts provided by Curtin does not tell her whether or why Curtin’s own Referential Index was different from the one she produced in any material ways. Curtin’s and Plaintiffs’ insistence that Kwon should simply run his scripts ignores the fact that doing so with no benchmark in the form of a contemporaneously prepared

²⁵ Importantly, Kwon’s opinion on replicability is not the only opinion she offers regarding the unreliability of Curtin’s method. She offers numerous other opinions regarding his failure to properly test other portions of his method and the administrative feasibility of his method. Thus, even if this Court were to exclude this portion of her report and testimony, the other critiques she offers of Curtin’s method are valuable and relevant as well.

Referential Index from Curtin renders Kwon unable to know whether Curtin's code or any of his underlying data inputs changed after he created his own Referential Index(es). Curtin's scripts are not the end point of step two of his method, as is indicated by the fact that at least two versions of his Referential Index exist. The fact that Curtin has not identified or produced the Referential Index that was the "final version" does not mean that Kwon's opinion that she was unable to replicate any of his Referential Indexes is irrelevant or unhelpful to this Court.

Curtin's and Plaintiffs' insistence that none of his Referential Indexes were final, and that all that Kwon needs to test his method are his scripts, are not sufficient reasons to exclude Kwon's opinion on replicability. Kwon did exactly what a rebuttal expert should do, which is to attempt to follow Curtin's instructions using his scripts and the same pool of data that he used to see if any of her results matched any of the versions of the Referential Index that Curtin created in his virtual environment. She did not misunderstand Curtin's report, and her opinions on replicability are highly relevant and useful to this Court. Thus, Plaintiffs' motion to exclude Kwon's opinions on the replicability of Curtin's Referential Index is denied.

4. Defendant's Discovery Responses

Plaintiffs next assert that Defendant's ability to create the Defendant Fact Sheet proves that a class is ascertainable. (ECF No. 293, pp. 27–28 (stating that "Blackbaud was able to determine which information was impacted for each named Plaintiff by querying [its own] databases using only [named Plaintiffs'] names and 'basic information' about them.")).²⁶ Defendant counters that

²⁶ Plaintiffs are referring to a two-step process that resulted in the creation of what the parties now refer to as the Defendant Fact Sheet. (ECF No. 342, pp. 35–36). Plaintiffs created "Plaintiff Fact Sheets" by gathering information from the named plaintiffs such as birth dates, email addresses, any Blackbaud customer to whom they believed that they had provided their information, and what data elements they believed had been exposed. *See, e.g.*, (ECF No. 333-1). These Plaintiff Fact Sheets were given to Defendant, and Defendant then searched for the named plaintiffs' information in the appropriate customer backup files by "rehydrating" those backup files, running queries to search for pieces of constituent data, and then inspecting various tables within the rehydrated backups to identify as much relevant information for a given constituent as possible. (ECF No. 294-17, p. 20). Through this process of searching the rehydrated backups using the information provided by the named plaintiffs, Defendant was able to produce Exhibit

its process of identifying the data elements belonging to the named plaintiffs in the course of producing the Defendant Fact Sheet was a manual and time-consuming process that was not designed to be used on a large scale and does not prove that an administratively feasible method of ascertaining the proposed classes or sub-classes exists. (ECF No. 329, p. 22; ECF No. 342 pp. 35–36). In the course of making this argument, Plaintiffs point to Curtin’s expert report, in addition to transcripts from the deposition of Jennifer Willson (the individual who appeared for Defendant at its 30(b)(6) deposition) and exhibits showing data gathered by Blackbaud (ECF Nos. 294-17, 294-29), as support for this proposition. Defendant has moved to exclude Curtin’s opinion that Defendant’s efforts to create its Defendant Fact Sheet “establish an administratively feasible method for identifying putative class members and their associated elements.” (ECF No. 342, pp. 35–36). This Court will address Plaintiffs’ argument that Defendant’s production of the Defendant Fact Sheet proves that a class can be ascertained without ruling on Defendant’s motion to exclude Curtin’s opinion on that point.²⁷ Plaintiffs could very likely advance this argument without Curtin’s opinion, as there is independent support for the argument found in Willson’s deposition transcripts and in the Defendant Fact Sheet itself, and this Court must therefore address it in any case.

i. Defendant Had the Plaintiff Fact Sheets

First, Plaintiffs’ argument overlooks the significance of the fact that Defendant was only able to effectively query the restored customer backup files because it had information from the

B to its Defendant’s Revised Fact Sheet. (ECF No. 329, p. 34 (explaining that “Blackbaud’s fact sheet . . . reflects the data stored for each Plaintiff by the customers Plaintiffs identified in their fact sheets”)).

²⁷ Although this Court is not ruling on Defendant’s motion to exclude Curtin’s opinion on Blackbaud’s creation of the Defendant Fact Sheet at this time, the Court will note that Curtin appears to be relying on deposition testimony and the Defendant Fact Sheet as the sole support for this opinion. (ECF No. 317-2, pp. 64–66). Curtin has not provided any additional analysis of Blackbaud’s work, tested any of Blackbaud’s scripts, or specifically compared its work to his method to support his opinion that the creation of the Defendant Fact Sheet demonstrates ascertainability. Further, he only spends four paragraphs of his 206-page report offering this opinion. Thus, it is doubtful that Curtin’s opinion on this matter would aid the fact finder.

named plaintiffs in its possession that it could use both as a starting point in crafting its queries and to validate the results of its queries. *See, e.g.*, (ECF Nos. 387-14, 387-15, 387-16, 387-17, 387-18). Plaintiffs have abandoned their argument that they can use email addresses as proxies for putative class members and that they can use those email addresses to contact putative class members. Thus, Plaintiffs no longer contend that they will be using the email addresses compiled in a theoretical Referential Index to begin compiling data profiles for each putative class member as was initially proposed. (ECF No. 380, pp. 6–7). Additionally, Plaintiffs have not tested, briefed, or otherwise demonstrated how they would collect information from putative plaintiffs to conduct a process similar to the process Defendant undertook in creating its Defendant Fact Sheet. They simply state that putative plaintiffs will “provide[]” it. (ECF No. 380, p. 7). However, implicit in this need for information from putative plaintiffs is the need to be able to reach those individuals in order to obtain their information. Plaintiffs have not tested, identified, or timely proposed a method of obtaining putative plaintiffs’ information, another fundamental flaw in this proposed method of ascertaining a class.

At the March 6-8 hearings, Plaintiffs proposed for the first time the creation of a website similar to the one used in the “*Equifax* data breach settlement” that could gather data from class members for use in searching the customer backup files to ascertain class members.²⁸ (ECF No. 496, pp. 29–30). This method has not been briefed or tested, and Plaintiffs are not permitted to make such an important and entirely new argument at oral argument. *See N. Carolina All. For Transp. Reform, Inc. v. U.S. Dep’t of Transp.*, 713 F.Supp. 2d 491, 510 (M.D.N.C. 2010) (“Raising . . . new arguments for the first time at oral argument undermines the purpose of orderly briefing

²⁸ For the same reasons stated below, *infra* Section III.A.6, arguments presented for the first time at a hearing can be properly disregarded by a court. Plaintiffs’ ever-evolving arguments regarding ascertainability and the methods by which a class could be ascertained in this case have made responding to the parties’ arguments and ruling on the parties’ motions a complicated task.

and risks subjecting an opponent to an unfair disadvantage.”). They have also suggested that additional Plaintiff Fact Sheets, or self-certified affidavits, similar to those created during discovery for the named plaintiffs could be created. This proposal has not been briefed or otherwise presented to this court outside of the March 6-8 hearings. (ECF No. 496, pp. 159–60). In addition to the fact that this proposal was not made in a timely fashion, this Court is persuaded by the same rationale that has guided other courts to hold that “[a]ffidavits from potential class members, standing alone, without ‘records to identify class members or a method to weed out unreliable affidavits,’ will not constitute a reliable and administratively feasible means of determining class membership.” *City Select Auto Sales Inc. v. BMW Bank of N. Am. Inc.*, 867 F.3d 434, 441 (3d Cir. 2017) (permitting use of affidavits to assist in ascertaining a class where there were other confirming records that could be obtained “with relative ease that would confirm [an individuals’] membership in the class”). Plaintiffs have not shown that confirming records could be used to “weed out unreliable affidavits” with “relative ease,” and this proposal was also not raised at an appropriate stage of this litigation, such that it is improper on both substantive and procedural grounds.²⁹

In any case, neither the *Equifax*-esque website proposal nor the self-certified affidavit proposal was made in a timely fashion, and this Court will not permit Plaintiffs to continually alter their ascertainability proposals in a manner that deprives Defendant of its ability to meaningfully defend itself. Further, even if Plaintiffs were to rely on such proposals, that would not eliminate

²⁹ Plaintiffs state in their opposition to Defendant’s motion to exclude Curtin’s report and testimony (ECF No. 380) that Curtin’s method will use “information provided by a putative class member.” They do not identify how this information will be obtained, and they do not raise this argument in their own Motion to Certify as a component of an independent method by which a class may be ascertained. The passing suggestion that information can be obtained from class members in order to ascertain a class buried within a responsive brief does not constitute raising that argument, and it certainly does not provide this Court or Defendant with sufficient detail to properly evaluate this suggestion by Plaintiffs. See *U.S. V. Dunkel*, 927 F.2d 955, 956 (7th Cir. 1991) (“Judges are not like pigs, hunting for truffles buried in briefs.”).

the need for validation of the resulting data profiles and the individuals to whom they belong to confirm that the data in question belongs to that person and has not been conflated or misidentified, which Curtin has acknowledged is a possible result of querying the customer backup files on a large scale. (ECF No. 494, p. 98).

Plaintiffs argued numerous times throughout their briefs and at the March 6–8 hearings that Defendant improperly conflates ascertainability and notice. (ECF No. 380, p. 7; ECF No. 494, pp. 72–73). The Court disagrees that the concepts were conflated but acknowledges that they are interrelated. The reason that Defendant and this Court must address Plaintiffs’ ability to reach putative class members at this stage of the litigation is because Plaintiffs’ proposed methods of ascertaining a class require information to be obtained from those class members to satisfy their proposed class definitions. *See* (ECF No. 293, pp. 12, 29 (stating that the third step in Curtin’s method is “to obtain information from putative Class and Subclass members” and noting that “basic information from claimants” will be needed to identify class members); ECF No. 496, p. 29 (stating that Plaintiffs would need to obtain “basic information” from class members in order to “tell them, you’re in the class or you’re not”)). This Court understands well the difference between ascertainability and notice, and this Court is not under the misimpression that Plaintiffs must “identify every class member at the time of certification.” *Krakauer v. Dish Network, L.L.C.*, 925 F.3d 643, 658 (4th Cir. 2019). However, when Plaintiffs themselves claim that information from class members is needed to ascertain a class, they cannot then insist that the Court and the opposing party are wrong to address whether Plaintiffs have timely and thoroughly demonstrated how that information would be obtained and whether a class can be ascertained using Plaintiffs’ proposed methods. *See EQT Prod. Co.*, 764 F.3d at 359 (noting that “[t]he fact that verifying ownership will

be necessary for class members to receive royalties does not mean it is not also a prerequisite to identifying the class”).

This leaves Plaintiffs with a glaring hole in all of their proposals when faced with the problem of how they would begin their process of searching for putative plaintiffs’ allegedly exposed data. They have no clear starting point now that they have abandoned their argument that Curtin can use the email addresses present in his Referential Index for this purpose. (ECF No. 380, pp. 7, 11). Defendant, on the other hand, had such a starting point—the Plaintiff Fact Sheets. This significant dissimilarity between Defendant’s method of searching for named plaintiffs’ data and Plaintiffs’ proposed method of searching for putative plaintiffs’ data undermines Plaintiffs’ argument that Defendant’s process demonstrates ascertainability. Defendant had information that Plaintiffs do not have as a starting point and that Plaintiffs have not demonstrated that they will have.

ii. Defendant’s Process Was Not Designed to be Automated or Scaled

Second, Defendant notes that its process of using the Plaintiff Fact Sheets to create its Defendant Fact Sheet was a manual and time-consuming process that was never intended to be scaled for use across 90,000 backup files and tens of millions of putative plaintiffs. (ECF No. 329, p. 22; ECF No. 342, pp. 35–36). Defendant has explained that it “[ran] searches for the data that it [] received from the plaintiffs against those rehydrated data files” and “manually reviewed all of the results . . . to . . . tailor the information . . . to the actual plaintiff that had been named in the case.” (ECF No. 496, p. 98). Defendant’s process was also clearly imperfect in light of Curtin’s discovery of information belonging to named plaintiff Philip Eisen in customer backup files that Mr. Eisen did not identify on his Plaintiff Fact Sheet, while Mr. Eisen’s information could not be found in backup files belonging to customers he identified as having his data. (ECF No. 342, pp.

35–36). Ultimately, Defendant conducted searches for specific kinds of data elements in the specific customer backup files that belonged to the customers that named plaintiffs self-identified as having their data. Defendant was only required to complete this task for thirty-four individuals. (ECF No. 333-22). Defendant’s ability to search a discrete set of customer backup files for specific data points belonging to thirty-four individuals using a manual process that took many hours to complete with the advantage of Plaintiff Fact Sheets to guide them tells this Court nothing about Plaintiffs’ ability to search 90,000 customer backup files associated with eleven different customizable products, in addition to the loose and unstructured files, for putative plaintiffs’ data. Plaintiffs need to undertake a materially different task than the one Defendant undertook in creating the Defendant Fact Sheet, rendering Defendant’s compilation of its Defendant Fact Sheet unhelpful to Plaintiffs in their efforts to demonstrate that a class is ascertainable.³⁰

Thus, Defendant’s ability to create the Defendant Fact Sheet does not demonstrate by a preponderance of the evidence that a class can be ascertained because (1) Defendant’s process was substantially manual and not made to be scaled and (2) Defendant had the Plaintiff Fact Sheets to use as a starting point in its process. *See E&G, Inc. v. Mount Vernon Mills, Inc.*, No. 6:17-CV-318-TMC, 2019 WL 4034951, at *3 (D.S.C. Aug. 22, 2019) (stating that a plaintiff must “show[] by a preponderance of the evidence that class certification is appropriate under Rule 23”) (citing *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 350–51 (2011)). Plaintiffs must undertake a broader and more complicated task, which would require them to present a viable method of obtaining data

³⁰ Plaintiffs also argue that Defendant used the “same process” to identify the thirty-four named plaintiffs’ information for which it searched and suggest that the similarity in Defendant’s process across the named plaintiffs demonstrates that there is an administratively feasible process that could be used to ascertain a class. (ECF No. 293, p. 28). Plaintiffs’ argument on this point is somewhat misleading because it disregards the fact that Blackbaud was able to use the same process repeatedly because it was searching for a discrete set of information provided to it by the named plaintiffs, and it was looking in a very limited set of customer backup files for that information. It is unsurprising that Defendant used the same process to look in a limited universe of customer backup files for a discrete set of information and confirm its presence. Plaintiffs’ task is much more open-ended.

from putative class members which Plaintiffs have said they will need, searching 90,000 customer backup files and loose files for individuals' data elements, and validating the output of those searches. Plaintiffs' task is far broader and more complex than the narrow task Defendant was asked to do, and Defendant's ability to create the Defendant Fact Sheet is not proof that Plaintiffs can undertake the larger task of ascertaining the proposed classes and sub-classes. *See EQT Prod. Co.*, 764 F.3d at 359–61 (noting that “[the named] plaintiffs [had] all been identified as gas estate owners” but still declining to find that the class as a whole could be ascertained due to the need for “complicated and individualized” review of land records in order to verify ownership).

5. Defendant's Notice Given to Customers

Thirdly, Plaintiffs contend that Defendant's ability to give notice to its customers of the fact of the breach and the possibility that certain kinds of constituent data were exposed indicates that a class can be ascertained in an administratively feasible manner. (ECF No. 293, pp. 28–29). Defendant counters that its efforts to notify *customers* of the breach never required it to identify individual *constituents* or actually view and validate their data. (ECF No. 329, p. 22). Curtin offered an opinion that “Blackbaud's own efforts to notify its customers demonstrate that it is feasible to identify class members using objective criteria” that Defendant has moved to exclude. (ECF No. 342, p. 37). Again, this Court will address Plaintiffs' argument regarding Defendant giving notice to its customers on its merits. Both Curtin's opinion and Plaintiffs' argument rest on testimony from Jennifer Willson and materials produced in discovery by Defendant showing its queries and their results that must be dealt with whether or not Curtin's opinion on this point is excluded. (ECF No. 317-2, pp. 67–71; ECF No. 293, pp. 28–29).

Plaintiffs' and Curtin's suggestion that Defendant's ability to notify its customers of the breach demonstrates that a class can be ascertained using objective criteria and in an

administratively feasible manner fails. Defendant points out that it never attempted to “identify constituents” who were affected by the breach, and it never viewed any constituent data to identify whether any of it was unencrypted PII. (ECF No. 342, p. 37). Instead, Defendant used queries “that would identify *potential* unencrypted Social Security Numbers, bank account numbers, and usernames and passwords” that were “stored in fields intended and/or fields that could reasonably be interpreted as being intended to store such information.”³¹ (ECF No. 342, p. 37). Once Defendant identified the potential presence of unencrypted PII in a customer’s database, it provided notice to those customers, who in turn were responsible to provide notice to individual constituents as each customer saw fit. (ECF No. 496, pp. 94–97). Importantly, Defendant’s customers did not necessarily take the step of identifying specific affected constituents, with several choosing to notify their entire constituent base of the breach. (ECF Nos. 333-26, 333-27 (stating that “[a]ll constituents in the database were notified” and that “[w]e were not able to identify the individuals that were impacted, so we [provided notice] to everyone.”)). Thus, the fact that individual constituents received notice does not indicate that those individual constituents can be or have been identified in an administratively feasible manner, either by Defendant or Defendant’s customers.³² Further, a representative of Defendant has testified at deposition that it searched “the ‘live’ databases at its BO3 data center in Boston” to provide notice to customers, not the affected customer backup files. (ECF No. 293, p. 28). In other words, the databases that

³¹ Defendant further emphasized during the Class Certification portion of the March 6-8 hearings that it merely “used a nine-digit pattern-matching stream” to find “potential instances where there would be nine-digit characters in a row.” (ECF No. 496, p. 95). However, Defendant did not validate these findings.

³² Defendant pointed out at the March 6-8 hearings that Plaintiffs’ notice argument is “overinclusive” because customers’ constituent notifications were not tailored to specific individuals, “underinclusive” because customers do not always have up-to-date contact information for constituents, and lastly that Plaintiff has provided no plan for how it would compel or obtain the cooperation of the 13,000 customers at issue. (ECF No. 496, pp. 95–96). For these reasons, in addition to the reasons listed above, including Plaintiffs’ failure to propose this plan in a timely manner, the Court is not persuaded by Plaintiffs’ argument.

Defendant searched to provide notice to customers were not the set of customer backup files that Plaintiffs must use to ascertain a class.

In summary, Defendant searched its set of live databases, not the 90,000 affected customer backup files, and performed pattern-matching that it did not validate in order to inform customers that their constituents' PII had *potentially* been impacted in the breach.³³ Those customers were not always able to provide individual notice to constituents, resulting in notice being given *en masse* to constituents, regardless of whether their data was impacted by the breach. This process of giving notice does not demonstrate that an administratively feasible method exists for ascertaining a class, particularly since none of the identified data was ever validated, and neither Defendant nor its customers took steps to identify specific individuals whose PII or PHI was affected.³⁴ Defendant's search of a different set of databases from the backup files Plaintiffs must use, its reliance on pattern-matching to show it possible PII that it did not need to validate, and its notice to customers without needing to identify affected individuals all differ materially from what Plaintiffs must do to demonstrate that the proposed classes and sub-classes are ascertainable. Plaintiffs must inspect the 90,000 customer backup files that were accessed during the breach, identify at least some data points for each individual and validate them, and they must identify individual class members, not merely customers, to satisfy their class definitions. The steps Defendant took to give notice to its customers are not comparable to the steps Plaintiffs would

³³ Plaintiffs have, at various times, implied that Defendant provided notice to individual constituents. (ECF No. 494, p. 56 (stating that "every single one of the class representatives . . . received a notice letter from Blackbaud.")). There is no evidence that Defendant ever directly provided notice to individual constituents or the named plaintiffs. Defendant merely provided notice to its **customers** of the breach, a fundamentally different task from the task of notifying individuals that specific unencrypted data elements were exposed.

³⁴ Further, Defendant never searched for evidence of affected PHI, rendering Plaintiffs' argument even less useful when considering the proposed state law sub-classes. (ECF No. 342, p. 37 (noting that Curtin acknowledged at his deposition that Blackbaud never searched for PHI)).

need to take to ascertain a class, and thus Defendant’s ability to give notice to its customers does not prove by a preponderance of the evidence that a class can be ascertained.

6. Defendant’s Use of Wirewheel

For the first time at the March 6–8 hearings, Plaintiffs turned to a singular paragraph in Curtin’s 206-page report that had not been highlighted or addressed in any of Plaintiffs’ prior briefs in this matter to argue that this paragraph presents a fourth method by which a class can be ascertained.³⁵ Although this Court is not required to consider this argument, since it was not raised in a motion or any other filings made by Plaintiffs, this Court will address it briefly. *See Synovus Bank v. Stevens Law Firm*, No. 4:19-CV-01411-SAL, 2020 WL 12788154, at *2 n.3 (D.S.C. Jul. 20, 2020) (declining to consider an unconscionability argument raised by a party because it was raised “for the first time at the hearing on the present motion”); *see also N. Carolina All. For Transp. Reform, Inc.*, 713 F.Supp. 2d at 510 (“Raising . . . new arguments for the first time at oral argument undermines the purpose of orderly briefing and risks subjecting an opponent to an unfair disadvantage.”). From the brief statements made by Plaintiffs’ and Defendant’s counsel at the class certification hearing, this Court is not persuaded by Plaintiffs’ argument. Defendant’s ability to

³⁵ Paragraph 153 in Curtin’s report states that:

“Blackbaud engineers provide California consumers who submit a CCPA Access request with a list of all their data elements that Blackbaud has sold in the past twelve months. In order to provide this information, Blackbaud uses a company called Wirewheel. . . . To use Wirewheel, a California resident uploads their driver’s license to the Wirewheel website to verify their identity, at which point their information is put into the Blackbaud engineering queue. . . . Blackbaud engineers then search their databases to identify the data Blackbaud has sold in the past twelve months pertinent to that consumer. . . . While this exercise pertains to Blackbaud’s cooperative database—the database Blackbaud uses to sell consumer data to third parties—it demonstrates the feasibility of identifying affected class members and the data elements affiliated with their personal information.”

(ECF No. 317-2, pp. 66–67).

This paragraph was not addressed in Plaintiffs’ Motion for Class Certification, its Reply, or its Response in Opposition to Defendant’s Motion to Exclude Curtin’s report and testimony on ascertainability. (ECF Nos. 293, 414, 380).

utilize a singular, live database that it maintains for the sole purpose of responding to CCPA requests does not in any way indicate that Defendant is necessarily able to restore and query 90,000 backup files of databases that were customized, maintained, and controlled by 13,000 separate customers. This apples-to-oranges comparison, which Plaintiffs waited until March 6, 2024 to raise for the first time, does not demonstrate that a class can be ascertained, and has nevertheless not been made in a timely manner.

7. Remaining Ascertainability Arguments

Having addressed Plaintiffs’ four proposed methods of ascertaining a class—Curtin’s method presented in his report, Defendant’s creation of its Defendant Fact Sheet, Defendant’s notice given to its customers, and Defendant’s use of Wirewheel to comply with the CCPA—the Court finds it necessary to address two ancillary arguments made by Plaintiffs. Additionally, the Court will briefly summarize its conclusions regarding why the proposed classes and sub-classes are not ascertainable in this case.

i. Use of Records Within a Defendant’s Control to Ascertain a Class

Importantly, Plaintiffs have noted that “courts do not look favorably upon the argument that records a defendant treats as accurate for business purposes are not accurate enough to define a class.” *Soutter v. Equifax Info. Servs., LLC*, 307 F.R.D. 183, 197–98 (E.D. Va. 2015); *see also In re Marriott*, 345 F.R.D. at 144–45. The Sixth Circuit, for instance, has stated that “the need to manually review files is not dispositive. . . . It is often the case that class action litigation grows out of systemic failures of administration, policy application, or records management that result in small monetary losses to large numbers of people. To allow that same systemic failure to defeat class certification would undermine the very purpose of class action remedies.” *Young v. Nationwide Mut. Ins. Co.*, 693 F.3d 532, 540 (6th Cir. 2012). Further, a district court in this circuit

has recently declined to credit a defendant's argument that its database was too unreliable to use to ascertain a class "because [defendant] used the NDS database to notify proposed class members of the breach." *In re Marriott*, 345 F.R.D. at 144–45.

This Court recognizes and agrees with the proposition that a defendant should not be able to hide behind problems of its own making with respect to records it maintains and controls. However, this case is markedly different from the cases Plaintiffs have highlighted on this point primarily because of the business-to-business relationship between Defendant and its customers and the fact that Defendant never used the data at issue in this case to provide notice directly to named plaintiffs.³⁶ Defendant is a Software-as-a-Service company, which means that its customers "control what data is collected, how it is stored, and where it is stored." (ECF No. 329, p. 13). Further, the data breach in this case was of Defendant's customer backup files, not a live environment. (ECF No. 329, p. 13). The 90,000 backup files that were accessed are not uniform as to any one customer or product because of the customers' ability to customize the products they purchase.³⁷ (ECF No. 329-3, pp. 56–57 (explaining that just among the products Curtin inspected, approximately 57% of those backup files had been customized)). Thus, the customer backup files at issue in this case are not merely data that Defendant has organized poorly and is now trying to argue cannot be used to ascertain a class. The data in question consists of customizable, varied

³⁶ The *Soutter* case involved Equifax, a well-known consumer reporting agency, that had gathered and organized data belonging to hundreds of millions of consumers. *Soutter*, 307 F.R.D. at 197. Because Equifax was able to provide the plaintiff in that case with "a list of the names and addresses of each consumer who made a dispute concerning a Virginia judgment on their file" within a specific time window, and because other record-keeping difficulties were either purely hypothetical or a product of Equifax's own unclear record-keeping, a class was found to be ascertainable in that case. *Id.* In *In re Marriott*, the defendant hotel chain is not a Software-as-a-Service company and has a direct relationship with the putative plaintiffs, whose data it collects and stores without any third party being involved or dictating how records are kept. *In re Marriott*, 341 F.R.D. at 138–39.

³⁷ Because Curtin only restored the most recent backup file(s) for each customer that he restored while crafting his method, this Court is not aware of the full extent of the variability that exists across all of the backup files for any one customer. (ECF No. 494, p. 90). Nonetheless, this Court knows such variability exists (ECF No. 494, p. 44), which makes this case less straightforward than the cases Plaintiffs have pointed to in suggesting that Defendant is improperly hiding behind records that it keeps and controls.

backup files that Defendant stores. However, Defendant did not dictate how those files were customized, whether that customization changed over time, what particular data was stored, and whether it was stored in encrypted or unencrypted form.

This case is therefore not comparable to *In re Marriott, Soutter*, or the numerous other cases Plaintiffs mentioned at oral argument. None of the cases addressed by Plaintiff, and certainly none that bind this Court, implicate the sort of three-party relationship between Defendant, its 13,000 customers, and their constituents that makes the proposed classes in this case so difficult to ascertain. Further, unlike in *Marriott*, Defendant never used its customer backup files to provide notice directly to class members, nor has it performed any other tasks using the data at issue (through its use of Wirewheel, its production of the Defendant Fact Sheet, or otherwise) that demonstrate that the proposed classes and sub-classes are ascertainable. As discussed above, Defendant notified its *customers* of the fact of the breach, and in some cases it used pattern-matching to notify customers that information that resembled social security numbers was present in the databases that were breached. However, Defendant never contacted any putative class members directly to provide them with notice, nor is there evidence that Defendant's customers did so. In numerous instances, Defendant's customers chose to provide notice to all constituents instead of attempting to identify which constituents' data was specifically affected. Thus, Defendant's observation that the customer backup files are widely varied, customized, and that Plaintiffs have not presented an administratively feasible way to ascertain a class is not a disingenuous argument but rather a factually correct statement that is supported by the record in this case.

ii. Whether Manual Review Can Preclude Ascertainability

Plaintiffs have further argued that this Court should not decline to find that the proposed classes and sub-classes are ascertainable simply because a large amount of data would need to be reviewed to satisfy the proposed class and sub-class definitions. (ECF No. 293, pp. 29–31). However, implicit in the ascertainability requirement is the reality that, at some point, the task of identifying class members can become too large and cumbersome for a court or a party to undertake. *See, e.g. Spotswood v. Hertz Corp.*, No. CV RDB-16-1200, 2019 WL 498822, at *6-7 (D. Md. Feb. 7, 2019) (stating that a “[p]laintiff cannot require [a] [d]efendant to manually search thousands of records to locate the putative class members” and that “the administrative difficulties involved with locating class members [manually] were too onerous”). Various cases within this circuit have addressed the extent to which manual review can preclude ascertainability, and there is not clear consensus regarding the point at which the amount of manual review precludes ascertainability. For instance, the District of Maryland in *Yates v. NewRez LLC* noted that a plaintiff is not required to prove that a class “can be perfectly generated at the touch of a button” in order to demonstrate ascertainability.³⁸ *Yates v. NewRez LLC*, No. CV TDC-21-3044, 2023 WL 5108803, at *5 (D. Md. Aug. 9, 2023). The *Kelly v. RealPage* court similarly observed that where plaintiffs have “identified the records they require, demonstrated they are in [the defendant’s] possession, and explained how those records can be used to verify putative subclass members,” a class is ascertainable. *Kelly v. RealPage Inc.*, 47 F.4th 202, 223–24 (3d Cir. 2022). Lastly, the

³⁸ The *Yates* court went on to say that “the need to review individual files” to ascertain class members is not a sufficient reason to deny class certification. (quoting *Byrd v. Aaron’s Inc.*, 784 F.3d 154, 171 (3d Cir. 2015)). However, the *Yates* court also observed that the need for individual review of loan documents in that case was limited to a narrow subset of the records at issue, which mitigated the problem of whether manual review would preclude ascertainability. *Id.* The need for individual review in this case extends to every putative plaintiff in the proposed classes and sub-classes.

Soutter case required manual review of records for just 1,000 consumers, which that court held to be a task that did not preclude ascertainability. *Soutter*, 307 F.R.D. at 197.

In this case, Plaintiffs have identified the records they require, and they have shown that they are in Defendant's possession, but as discussed above, they have not provided a timely raised, sufficiently tested, thoroughly briefed, and administratively feasible method of using those records to ascertain a class. Additionally, this Court has considered the Fourth Circuit's recent guidance in *Career Counseling* and determined that the Fourth Circuit is instructing this Court and others that a certain degree of manual review can preclude ascertainability. *See Career Counseling*, 91 F.4th 202, at 211–12 (affirming the district court's ruling that the need to make "an individualized inquiry as to whether each [fax] recipient was using a stand-alone fax machine at the relevant time" for "more than 20,000 recipients" precluded ascertainability). In this case, far more than 20,000 individual inquiries would need to be made to ascertain the proposed classes and sub-classes. For the same reasons that the district court in *Career Counseling* court held that 20,000 inquiries into whether putative plaintiffs had a certain kind of fax machine precluded ascertainability, this Court finds that the hundreds of millions of inquiries that would be required to determine whether a putative plaintiff's data is located in the 90,000 customer backup files at issue and to validate those data elements place this case far outside the bounds of the "administrative feasibility" requirement imposed by this circuit.

Plaintiffs have suggested, albeit without sufficient support, that methods exist whereby the customer backup files can be restored and queried in an automated manner. Their assertion has not been demonstrated by a preponderance of the evidence. Curtin did not test or otherwise demonstrate an ability to do so within an acceptable rate of error, Defendant has not done so on a scale that proves to this Court that it can be done in a feasible manner, and Plaintiffs have offered

nothing further aside from bald assertions that the customer backup files can be queried and the putative plaintiffs can be identified. Further, even if Plaintiffs were to use automated processes, Plaintiffs have not shown that their proposed class and sub-class definitions can be met without the need for manual intervention and validation in order to confirm that a putative plaintiff's data has not been improperly conflated with data belonging to someone else.

Plaintiffs agree that they must identify (1) whether someone's unencrypted data was exposed in the breach and (2) what specific elements were exposed. (ECF No. 414, p. 7). Plaintiffs also acknowledge that putative plaintiffs' states of residence must be determined, although they contend that doing so is a "claims administration" problem rather than an ascertainability problem. (ECF No. 414, pp. 11–12). All of that information is useless unless it is confirmed that the individual in question was in fact a Blackbaud customer at the relevant time and that the data elements being attributed to them are correct and do in fact belong to them. Plaintiffs suggested at the March 6-8 hearings that all this Court needs to do is undertake a simple "data in, data out" inquiry. (ECF No. 494, p. 72). This Court does not agree. A profile of non-validated data elements does not tell this Court whether a person is, in fact, a member of the class, especially in light of the possibility of "mixing and matching" of data elements across individuals that Plaintiffs have acknowledged. (ECF No. 494, p. 98). More than a binary inquiry is required to determine whether a putative plaintiff has been correctly identified (whether that is done using an email address, driver's license number, or some other primary key) and whether any of the data elements attributed to them do, in fact, belong to them. Given the estimated size of the class and the breadth of the customer backup files and the loose files, even a limited amount of manual intervention under any of the methods Plaintiffs have proposed would quickly become "too onerous" for the Court and the parties to undertake. *See Spotswood*, 2019 WL 498822, at *6.

iii. Summary of Ascertainability Conclusions

In sum, “class certification is inappropriate when ‘class members are impossible to identify without extensive and individualized fact-finding.’” *Career Counseling*, 2021 WL 3022677, at *12. As discussed above, each method of ascertaining a class that Plaintiffs have proposed is flawed. Each would also require this Court to engage in significant individualized fact-finding if implemented. The inquiries that must be undertaken to ascertain class members in this case go far beyond a simple determination of whether Plaintiffs can run searches that will return results. Millions of individualized inquiries would be required to determine, *at minimum*, (1) whether the putative plaintiffs are in fact members of the class, in light of the clear risk of misidentification or conflation of email addresses and other identifiers that Plaintiffs have acknowledged, (2) which unencrypted data elements belonging to each putative plaintiff were exposed, due to the need for validation of each putative plaintiffs’ data profile to ensure that the proposed class and sub-class definitions are met and that data is not being mis-attributed to the wrong individual, and (3) which state a putative sub-class member resided in at the time of the breach. Further, none of Plaintiffs’ proposed methods would relieve this Court of the burden of needing to verify the aforementioned information, as all proposed methods involve an uncertain risk of error, the unavoidable need for manual intervention and validation, and therefore an untenable degree of individualized inquiry for each putative plaintiff. With an estimated class size of up to 1.5 billion individuals, Plaintiffs’ proposed classes and sub-classes cannot be ascertained without significant individualized inquiry at a scale that is not administratively feasible for Plaintiffs, this Court, Defendant, or any individuals or entities acting at their direction to undertake. Therefore, Plaintiffs have not met their burden of demonstrating that class certification is appropriate in this case.

This conclusion is bolstered by the fact that courts in this circuit have declined to ascertain classes involving less labor-intensive review and fewer individualized determinations than this case would require. Recently, the Fourth Circuit in *Career Counseling* affirmed the district court’s determination that 20,989 individual determinations of whether “a fax number . . . was linked to a stand-alone fax machine” on the relevant date would require “individualized inquiry” that rendered the class not ascertainable. *See Career Counseling*, 2021 WL 3022677, at *11–12. Just as the *Career Counseling* court determined that more than 20,000 individualized determinations regarding putative plaintiffs’ use of a stand-alone fax machine precluded ascertainability, tens of millions of individualized determinations regarding a putative plaintiff’s presence in the customer backup files, the accuracy of a list of their specific exposed data elements, and their state residency make these proposed classes and subclasses far from ascertainable.

Further, the cases Plaintiffs primarily rely upon to argue that ascertainability should be found are either not controlling or are distinguishable. The *Soutter* case involved a class of approximately 1,000 people, with the court estimating that a few attorneys could review the necessary documents “in a matter of days.” *Soutter*, 307 F.R.D. at 197. The *Kelly v. RealPage* case from the Third Circuit involved a review of two databases controlled and organized by the defendant, without any third party customer’s involvement. *See Kelly*, 47 F.4th at 202. Curtin, on the other hand, has acknowledged that any process of searching the 90,000 customer backup files and loose files that resembles the process he proposed would likely require thousands of hours of work to complete. (ECF No. 494, p. 100). Lastly, none of the cases Plaintiffs address required information to be gathered from putative plaintiffs at such a large scale for a class to be ascertained in the first place. This case presents a far more onerous degree of review, both in terms of the

number of proposed class members and the amount of data that must be gathered and queried, than any of the cases Plaintiffs point to in arguing that a class can be ascertained in this case.³⁹

In sum, none of the methods Plaintiffs have proposed for ascertaining a class are administratively feasible on their face. Further, none of Plaintiffs' proposed methods have eliminated the need for this Court, Defendant, or some other party to make tens of millions of individualized inquiries in order to determine whether a given individual satisfies Plaintiffs' proposed class and sub-class definitions.⁴⁰ Very few courts in this country thus far have been willing to certify Rule 23(b)(3) classes "involving individual consumers complaining of a data breach." *In re Marriot*, 341 F.R.D. at 172. Given Plaintiffs' failure to provide this Court with an administratively feasible method of ascertaining class members, this Court declines to join the minority of courts that have certified a class in a consumer data breach case such as this.

8. Rule 23's Other Requirements

Rule 23(a) requires a class to satisfy the four requirements of "numerosity," "typicality," "commonality," and "adequacy" in order for a class to be certified. Fed. R. Civ. P. 23(a). Rule 23(b)(3) further requires that any common questions of law or fact must "predominate over any questions affecting only individual members." Fed. R. Civ. P. 23(b)(3). Lastly, Rule 23(b)(2) requires that "the party opposing the class [have] acted or refused to act on grounds that apply generally to the class, so that final injunctive relief . . . is appropriate respecting the class as a whole." Fed. R. Civ. P. 23(b)(2). Because this Court's decision ultimately rests on its conclusion

³⁹ It is also worth noting that no court in the Fourth Circuit has dealt with a data breach involving a class this large, especially not one that involved a business-to-business defendant. As Defendant correctly noted at the March 6-8 hearings, this case places this Court in "uncharted . . . territory." Further, much of the authority Plaintiffs have leaned upon most heavily in making their ascertainability arguments comes from the Third Circuit, which does not bind this court, although it can be instructive.

⁴⁰ As addressed in greater detail above, these inquiries would be necessary in light of the risk of misidentification or conflation of email addresses and other data elements across putative plaintiffs that Plaintiffs have recognized, the fact that data storage practices are not consistent across customers, and the need to verify state residency in light of putative plaintiffs' possible change in state of residency, at minimum.

that Plaintiffs have not met their burden of proving that their proposed classes and sub-classes can be ascertained, discussion of the remaining elements of 23(a) and 23(b) is unnecessary. However, many of the issues affecting ascertainability—namely the variability in data storage practices across Defendant’s customer base, the differences in the kinds of data stored for each putative plaintiff, the differences in the functions served by each Blackbaud customer, and the variability in the putative plaintiffs’ own circumstances such as prior exposures of the same data at issue in this case—cast doubt as to whether Plaintiffs could properly satisfy the requirements of commonality, typicality, and predominance.⁴¹

B. The Parties’ Other Pending *Daubert* Motions

Plaintiffs and Defendant have each moved to exclude all of the other party’s *Daubert* experts on various grounds. These motions are denied as moot, in light of the Court’s ruling that the proposed classes and sub-classes are not ascertainable.

C. The Effect of Declining to Certify a Class

This Court recognizes that “there remains the problem of how to deal with conduct that inflicts small amounts of damage on large numbers of people.” *In re Asacol Antitrust Litig.*, 907 F.3d 42, 56 (1st Cir. 2018). The First Circuit has recently aptly expressed many sentiments that this Court shares regarding the purpose of class action litigation and the decision not to certify a class even in the face of allegations that, if proven, were surely harmful and avoidable. The First Circuit, in *In re Asacol Antitrust Litig.*, noted that:

⁴¹ This Court is also concerned about Plaintiffs’ ability to prove and calculate damages as they proposed in their Motion to Certify. (ECF No. 293, pp. 36–43). Specifically, Plaintiffs may be unable to place putative plaintiffs into tranches of risk and calculate which data breach response packages they should receive, given that Plaintiffs’ proposed methods of doing these two things are predicated upon Curtin’s ability to identify the specific exposed data elements belonging to a given putative plaintiff. (ECF No. 293, pp. 40–41). Plaintiffs’ other proposed damages model, which centers around alleged loss in value of plaintiff data, would also require Plaintiffs to identify specific data elements in order to calculate their value. (ECF No. 293, pp. 42–43). This Court has determined that Curtin’s method is unreliable and unhelpful under Rule 702 and *Daubert*, leaving Plaintiffs’ ability to calculate damages as originally proposed in uncertain territory.

Rule 23 serves as an important tool to address many such situations. *See Mace v. Van Ru Credit Corp.*, 109 F.3d 338, 344 (7th Cir. 1997) (“The policy at the very core of the class action mechanism is to overcome the problem that small recoveries do not provide the incentive for any individual to bring a solo action.”); *Castano v. Am. Tobacco Co.*, 84 F.3d 734, 748 (5th Cir. 1996) (noting that “negative value” suits provide the “most compelling rationale for finding superiority in a class action”). But that fact grants us no license to create a Rule 23(b)(3) class in every negative value case by either altering or reallocating substantive claims or departing from the rules of evidence. Moreover, there are other tools available to address the problem of low-value, high-volume claims that pose individual issues of causation. Regulators may sue, *see, e.g., FTC v. Actavis, Inc.*, 570 U.S. 136, 141 (2013); governments may bring parens patriae claims, *see, e.g., New Hampshire v. Purdue Pharma*, No. 17-cv-427, 2018 WL 333824, at *1 (D.N.H. Jan. 9, 2018); substantive laws may provide presumptions available to all class members, *see, e.g., Halliburton*, 134 S.Ct. at 2411-12; and private lawyers may marshal the threats of res judicata and fee shifting to induce aggregate settlements when liability is clear.

In re Asacol Antitrust Litig., 907 F.3d 42, 56 (1st Cir. 2018).

In this case, several of these alternative safeguards have been implemented. Forty-nine of the fifty state attorneys general have entered into a significant settlement with Defendant as a result of this data breach. Those settlements include a requirement that Defendant improve its use of firewalls, intrusion detection, and dark web monitoring, in addition to improving its incident and breach response plans. (ECF No. 496, p. 71). The SEC has entered a Cease-and-Desist Order instructing Defendant to cease and desist from enumerated violations of the Securities Act and the Exchange Act and to pay a fine. (ECF No. 319-1). The FTC has issued a “Decision and Order” that enjoins Defendant from many of the data security practices that Plaintiffs allege caused and then misrepresented the nature of the data breach. (ECF No. 480-3). That Order specifically requires Defendant to adhere to specific data deletion and data detention procedures and to obtain information security assessments from a third party periodically for twenty years. *Id.*

Plaintiffs argue that declining to certify a class in this case signals that certain defendants cannot be held accountable if they are big enough and if they cause significant enough harm. This Court disagrees. Declining to certify the proposed classes and sub-classes in this case signals that

the requirements put in place by the Fourth Circuit and the Federal Rules of Civil Procedure are not mere boxes to be checked. The Court must emphasize that this decision does not leave Defendant free to carry on with reckless abandon. Nor does it forestall the prospect of relief by individuals actually harmed by the allegations in this action. Individual plaintiffs who believe that their data has been exposed and that they have been harmed as a result are free to bring suit individually. *See, e.g., Thorn v. Jefferson-Pilot Life Ins. Co.*, 445 F.3d 311, 318 (4th Cir. 2006) (stating that “[t]he class-action device is the exception to the strong default rule that ‘a party in federal court may vindicate only his own interests’”). Defendant will have to reckon with any individual lawsuits that are brought on their merits. To be sure, this MDL already contains over twenty tag-along actions consisting of claims brought by constituents or customers of Defendant.

Finally, this Court’s decision to decline to certify the proposed classes and sub-classes should not be taken as a stamp of approval of Defendant’s data security practices and its response to the data breach. If Plaintiffs’ allegations are true, Defendant could have taken greater precautions to protect customers and their constituents and could have correctly represented the extent of the breach upon initially discovering that it had occurred. (ECF No. 496, pp. 4–12). Nevertheless, the appropriate mechanism in this case for pursuing Defendant civilly for any damages incurred because of the breach and Defendant’s response to it is not the class action Plaintiffs’ counsel has proposed.

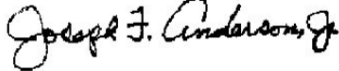
IV. CONCLUSION

Plaintiffs’ Motion to Certify a Class (ECF No. 292) is denied because Plaintiffs have failed to demonstrate that the proposed classes and sub-classes are ascertainable. Plaintiffs’ Motion to Exclude the Report and Testimony of Sonya Kwon (ECF No. 419) is denied, Defendant’s Motion

to Exclude the Report and Testimony of C. Matthew Curtin (ECF No. 341) is granted in part, and all other pending *Daubert* motions are denied as moot.

IT IS SO ORDERED.

May 14, 2024
Columbia, South Carolina


Joseph F. Anderson, Jr.
United States District Judge